HONI SOIT QUI MAL Y PENSE
DIEU ET MON DROIT

AGRICULTURAL RESEARCH INSTITUTE
PUSA

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A—MATHEMATICAL AND PHYSICAL SCIENCES

VOL CLXIV

18879
IARI

LONDON
Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N W 1

18 February 1938

18879

PRINTED IN GREAT BRITAIN BY
WALTER LEWIS, M A
AT THE CAMBRIDGE UNIVERSITY PRESS

CONTENTS

SERIES A VOL CLXIV

No A 916—7 January 1938

PAGE

The seiches in a strait connecting two seas By G R Goldsbrough, F R S 1

Production and dissipation of vorticity in a turbulent fluid By G I Taylor, F R S 15

Galvano magnetic effects in bismuth alloys By N Thompson 24

Hyperfine structure and nuclear moments of aluminium By D A Jackson and H Kuhn (Plate 1) 48

On anomalous vibrational spectra By M Blackman 62

Crystal growth from solutions By W F Berg (Plate 2) 79

Studies of region *E* of the ionosphere By J E Best F T Farmer and J A Ratcliffe 96

Stability of polyatomic molecules in degenerate electronic states II Spin degeneracy By H A Jahn 117

Progressive lightning IV The discharge mechanism By B F J Schonland 132

No A 917—21 January 1938

The theory of the photolysis of silver bromide and the photographic latent image By R W Gurney and N F Mott, F R S 151

Self consistent field with exchange for calcium By D R Hartree, F R S and W Hartree 167

On the statistical theory of isotropic turbulence By T de Kármán and L Howarth 192

The effect of temperature on the photochemical bleaching of visual purple solutions By H J A Dartnall, C F Goodeve and R J Lythgoe 216

Instability of fluids heated from below By K Chandra (Plates 3–6) 231

PAGE

The production of neutrons by bombardment of beryllium with α particles By T Bjerge 243

On the penetrating component of cosmic radiation By H J Bhabha 257

The hydrolysis of the methyl halides By E A Moelwyn Hughes 295

No A 918—4 February 1938

Significance tests for continuous departures from suggested distributions of chance By H Jeffreys, F R S 307

New bands ending on the $1s\sigma 2p\sigma\,^1\Sigma_u$ state of H_2 By O W Richardson, F R S 316

The two dimensional hydrodynamical theory of moving aerofoils II By Rosa M Morris 346

The crystal structure of pentaerythritol tetracetate By T H Goodwin and R Hardy 369

The electronic structure of some polyenes and aromatic molecules IV The nature of the links of certain free radicals By C A Coulson 383

The electronic structure of some polyenes and aromatic molecules V A comparison of molecular orbital and valence bond methods By G W Wheland 397

The electronic structure of some polyenes and aromatic molecules VI Phenylethylene, stilbene, tolane and the phenylmethyl radical By W G Penney and G J Kynch 409

The rotational energy levels of a diatomic molecule in a tetrahedral field By H M Cundy 420

X ray analysis of the dibenzyl series V Tolane and the triple bond By J Monteath Robertson and I Woodward 435

No A 919—18 February 1938

Relaxation methods applied to engineering problems II Basic theory, with applications to surveying and to electrical networks, and an extension to gyrostatic systems By A N Black and R V Southwell, F R S 447

PAGE

The reflexion coefficients of ionospheric regions By E V Appleton, F R S and J H Piddington 467

The spectrum of turbulence By G I Taylor F R S 476

Structure of stretched rubber By C J B Clews and F Schoszberger (Plate 7) 491

Statistical mechanics of the adsorption of gases at solid surfaces By F J Wilkins 496

The adsorption of argon nitrogen and oxygen on smooth platinum foil at low temperatures and pressures By F J Wilkins 510

The emission band spectrum of chlorine (Cl^{+}) II By A Elliott and W H B Cameron (Plate 8) 531

On the solution of the laminar boundary layer equations By L Howarth 547

The crystal structure of insulin I The investigation of air dried insulin crystals By Dorothy Crowfoot (Plate 9) 580

Index 603

The Seiches in a Strait Connecting Two Seas

By G R Goldsbrough, F R S

(*Received* 29 *June* 1937)

1—Introduction

The seiches or free long waves in lakes or seas with complete boundaries have been determined in many cases When the boundaries are incomplete the problem presents greater difficulties and the results are less definite Experimental and theoretical studies have been made by many investigators (Harris 1908, Honda and others 1908)

The analysis in elliptic co ordinates used by the writer (1930 and 1936) for certain problems of completely bounded seas lends itself somewhat readily to the problem of a strait connecting two open seas For a certain law of depth the general equations are reducible and solutions follow without much difficulty

For exact solutions it is necessary to solve the period equation by continued approximation But it is shown that in an important group of cases very approximate results can be written down explicitly and interpretation then follows readily

2—The equations for a strait joining two seas

It has been shown by the writer (1930) that the equations of motion of long waves of period $2\pi/\lambda$ in elliptic co-ordinates can be reduced to the equation

$$\frac{\partial}{\partial\xi}\left(h\frac{\partial\zeta}{\partial\xi}\right)+\frac{\partial}{\partial\eta}\left(h\frac{\partial\zeta}{\partial\eta}\right)+H^2\lambda^2\zeta/g = 0, \tag{1}$$

where $H = c^2(\cosh^2\xi-\cos^2\eta)$, h is the depth, and ζ is the displacement of the surface at the point (ξ, η)

Take $$h = h_0(\cosh^2\xi - a)(a-\cos^2\eta), \quad 0<a<1 \tag{2}$$

In the problem as defined by (1), h must be positive or zero Hence $0\leqslant\cos^2\eta\leqslant a$ The surface of the corresponding sea lies between the two branches of the hyperbola $\cos^2\eta = a$, and the bed of the sea has the form of a hyperbolic paraboloid

The maximum depth on the transverse axis of the hyperbola is at the centre, and the depth increases indefinitely with increase of distance from this point along the conjugate axis The equation (1) is based upon the assumption that the depth is everywhere small compared with the wavelength At points along the conjugate axis sufficiently remote, this condition clearly will be broken As will be seen later, however, the wave motions diminish in amplitude with recession from the transverse axis, and at points where the equations of long waves cease to apply the motion is practically zero

On substituting (2) in (1) we have

$$(a-\cos^2\eta)\frac{\partial}{\partial\xi}(\cosh^2\xi-a)\frac{\partial\zeta}{\partial\xi}+(\cosh^2\xi-a)\frac{\partial}{\partial\eta}(a-\cos^2\eta)\frac{\partial\zeta}{\partial\eta}$$
$$+\kappa^2(\cosh^2\xi-\cos^2\eta)\,\zeta=0,$$

where $\kappa^2=\lambda^2c^2/gh_0$

The variables are separable by putting $\xi = M(\xi)\,N(\eta)$ Then we have

$$\left.\begin{aligned}\frac{d}{d\xi}(a-\cosh^2\xi)\frac{dM}{d\xi}-\kappa^2M+k(a-\cosh^2\xi)\,M&=0,\\ \frac{d}{d\eta}(a-\cos^2\eta)\frac{dN}{d\eta}+\kappa^2N-k(a-\cos^2\eta)\,N&=0,\end{aligned}\right\}\tag{3}$$

k being an undetermined constant

Substitute $$k=-4\sigma(\sigma-1),$$
$$\kappa^2=4\sigma(\sigma-1)(q-a),$$

and these equations reduce to

$$\left.\begin{aligned}\frac{d}{d\xi}(a-\cosh^2\xi)\frac{dM}{d\xi}-4\sigma(\sigma-1)(q-\cosh^2\xi)\,M&=0,\\ \frac{d}{d\eta}(a-\cos^2\eta)\frac{dN}{d\eta}+4\sigma(\sigma-1)(q-\cos^2\eta)\,N&=0\end{aligned}\right\}\tag{4}$$

These are the same in form as those used by the writer (1936, p 13) and may be treated in a similar manner

The polynomial solutions of these equations hitherto used are, however, inappropriate to the present problems We therefore seek solutions in the form of infinite series

Put $\nu = \cosh^2\xi$, $\mu = \cos^2\eta$ and equations (4) become

$$\nu(\nu-1)(\nu-a)\frac{d^2M}{d\nu^2} + (2\nu^2 - \tfrac{3}{2}\nu - a\nu + \tfrac{1}{2}a)\frac{dM}{d\nu} + \sigma(\sigma-1)(q-\nu)M = 0, \qquad (5\,\text{i})$$

$$\mu(\mu-1)(\mu-a)\frac{d^2N}{d\mu^2} + (2\mu^2 - \tfrac{3}{2}\mu - a\mu + \tfrac{1}{2}a)\frac{dN}{d\mu} + \sigma(\sigma-1)(q-\mu)N = 0 \qquad (5\,\text{ii})$$

The solutions of these equations must satisfy two conditions

(*a*) N must be finite when $\mu = a$,

(*b*) $M \to 0$ as $\nu \to \infty$.

For the second condition a solution of (5 i) in powers of $1/\nu$ is required. On putting $\theta = 1/\nu$, equation (5 i) becomes

$$\theta(1-a\theta)(1-\theta)\frac{d^2M}{d\theta^2} + (\tfrac{3}{2}a\theta^2 - \tfrac{1}{2}\theta - a\theta)\frac{dM}{d\theta} + \sigma(\sigma-1)\left(q - \frac{1}{\theta}\right)M = 0 \qquad (6)$$

The point $\theta = 0$ is a regular singularity. The solution in the vicinity of this point is $M = \Sigma a_n \theta^{c+n}$, where the index c is σ or $1-\sigma$, and, for the first index,

$$a_{n+1}(n+1)(2\sigma+n) + a_n\{\sigma(\sigma-1)q - (\sigma+n)(\sigma+n-\tfrac{1}{2}) - a(\sigma+n)^2\}$$
$$+ a_{n-1}\, a(\sigma+n-1)(\sigma+n-\tfrac{1}{2}) = 0 \qquad (7)$$

On writing $N_{n+1} = a_{n+1}/a_n$, this relation becomes

$$N_{n+1}(n+1)(2\sigma+n) + \{\sigma(\sigma-1)q - (\sigma+n)(\sigma+n-\tfrac{1}{2}) - a(\sigma+n)^2\}$$
$$+ a(\sigma+n-1)(\sigma+n-\tfrac{1}{2})/N_n = 0 \qquad (8)$$

For large values of n, (8) has the limiting form

$$N_{n+1} - (1+a) + a/N_n = 0, \qquad (9)$$

which is satisfied by $N_n = 1, a$, for all n.

More exactly we find in the first alternative, $N_n \sim 1 - \frac{3}{2n}$. For large values of n the coefficients a_n then approximate to those of the expansion of $(1-\theta)^{\frac{1}{2}}$, and M has the form $\Phi + (1-\theta)^{\frac{1}{2}}\Psi$. By the argument of Lamb (1932) this would produce a discontinuity in the ξ component of the velocity on the transverse axis.

The series solution of (6) must be convergent for $\theta = 1$. Hence, since $a < 1$, the quantities σ, q must be determined so that while (8) is satisfied $N_n \to a$.

The equation (5 ii) has a solution for the regular singularity $\mu = 0$ in the form $\Sigma a'_n \mu^{c+n}$, where $c = 0$ or $\frac{1}{2}$, and

$$a'_{n+1}\; a(n+1)(n+\tfrac{1}{2}) + a'_n\{\sigma(\sigma-1)q - n^2(1+a) - \tfrac{1}{2}n\}$$
$$+ a'_{n-1}(n-\sigma)(n+\sigma-1) = 0, \quad (10)$$

for the first index, and

$$a'_{n+1}\; a(n+1)(n+\tfrac{3}{2}) + a'_n\{\sigma(\sigma-1)q - (n+\tfrac{1}{2})^2(1+a) - \tfrac{1}{2}(n+\tfrac{1}{2})\}$$
$$+ a'_{n-1}\{n^2 - \tfrac{1}{4} - \sigma(\sigma-1)\} = 0, \quad (11)$$

for the second index

Consider the solution defined by (10) If σ is integral the series terminates and q is determined by an algebraic equation Since σ must also satisfy (8) it cannot in general be integral, and the series of (10) will be infinite

Putting $N'_{n+1} = a'_{n+1}/a'_n$, the recurrence formula (10) reduces to

$$N_{n+1}\; a(n+1)(n+\tfrac{1}{2}) + \{\sigma(\sigma-1)q - n^2(1+a) - \tfrac{1}{2}n\}$$
$$+ (n-\sigma)(n+\sigma-1)/N'_n = 0 \quad (12)$$

For large values of n, (12) has the limiting form

$$N'_{n+1}\; a - (1+a) + 1/N'_n = 0, \quad (13)$$

which is satisfied by $N'_n = 1$, $1/a$ for all n, or more exactly in the second case $N'_n \sim 1/a(1-1/n)$ Since in the problem formulated $0 \leqslant \mu \leqslant a < 1$, the constants σ, q must be determined so that while (12) is satisfied, $N'_n \to 1$

A similar conclusion holds good for the series defined by (11)

The formulae (7), (10), (11) lead in the usual manner to infinite continued fractional equations The problem is now to determine σ, q from (7) and (10), or from (7) and (11) with in each case the appropriate convergence conditions

The equation (6) has a second index $1-\sigma$ associated with the point $\theta = 0$, which has so far not been used Inspection shows that if (10) is satisfied by a certain value of σ it is also satisfied when the same value is placed for $1-\sigma$ Hence if in (7) we replace σ by $1-\sigma$ and use it in conjunction with (10), precisely the same solutions will appear as before

In (8) put $p = \sigma(\sigma-1)q$,

$$L_n = \frac{p - (\sigma+n)(\sigma+n-\frac{1}{2}) - a(\sigma+n)^2}{(n+1)(n+2\sigma)},$$

$$y_n = \frac{a(\sigma+n-1)(\sigma+n-\frac{1}{2})}{(n+1)(n+2\sigma)},$$

then
$$\left.\begin{aligned} N_{n+1}+L_n+y_n/N_n &= 0, \\ N_n &= -\frac{y_n}{L_n+N_{n+1}} \end{aligned}\right\} \tag{14}$$
and

Similarly, in (12) put

$$L'_n = \frac{p-n^2(1+a)-\frac{1}{2}n}{a(n+1)(n+\frac{1}{2})},$$

$$y'_n = \frac{(n-\sigma)(n+\sigma-1)}{a(n+1)(n+\frac{1}{2})},$$

then
$$\left.\begin{aligned} N'_{n+1}+L'_n+y'_n/N'_n &= 0, \\ N'_n &= -\frac{y'_n}{L'_n+N'_{n+1}} \end{aligned}\right\} \tag{15}$$
and

These forms will be found useful in the later calculations

3—A FURTHER TYPE OF SOLUTION

The singularity $\nu = 1$ of equation (5 i) appears in the range of the variable of the present problem But, as $0 \leqslant \mu \leqslant a < 1$, the singularity $\mu = 1$ does not appear in (5 ii)

On substituting $M = \sqrt{(\nu-1)}\,w$, equation (5 i) reduces to

$$\nu(\nu-1)(\nu-a)\frac{d^2w}{d\nu^2}+(3\nu^2-2a\nu-\tfrac{3}{2}\nu+\tfrac{1}{2}a)\frac{dw}{d\nu}$$
$$+\{\tfrac{3}{4}\nu-\tfrac{1}{4}a+\sigma(\sigma-1)(q-\nu)\}w = 0 \tag{16}$$

On putting $\nu = 1/\theta$, this reduces to

$$\theta(1-\theta)(1-a\theta)\frac{d^2w}{d\theta^2}+(\tfrac{3}{2}a\theta^2-\tfrac{1}{2}\theta-1)\frac{dw}{d\theta}$$
$$+\left\{\frac{3}{4}\,\frac{1}{\theta}-\frac{1}{4}a+\sigma(\sigma-1)\left(q-\frac{1}{\theta}\right)\right\}w = 0 \tag{17}$$

For a solution valid in the vicinity of $\theta = 0$, put $w = \Sigma a_n \theta^{c+n}$ The indices are $\sigma+\frac{1}{2}$ and $\frac{3}{2}-\sigma$, and associated with the first is the recurrence formula

$$a_{n+1}\{(\sigma+n+\tfrac{3}{2})(\sigma+n-\tfrac{1}{2})+\tfrac{3}{4}-\sigma(\sigma-1)\}$$
$$+a_n\{\sigma(\sigma-1)q-(\sigma+n)(\sigma+n+\tfrac{1}{2})-a(\sigma+n)^2\}$$
$$+a_{n-1}\{a(\sigma+n)(\sigma+n-\tfrac{1}{2})\} = 0 \tag{18}$$

When (18) has been solved as before with the appropriate conditions, the solution of (5 i) takes the form

$$M = \sqrt{(\nu-1)}\,\nu^{-\sigma-\frac{1}{2}}\Sigma a_n \nu^{-n}$$

Hence for a valid type of motion, $\sigma > 0$

As before the solution associated with the index $\frac{3}{2}-\sigma$ introduces nothing new

Collecting the results we see that there are four types of solution of the equation (1) which apply to the problem as at present formulated

(i) the association of (7) with (10),

(ii) the association of (7) with (11),

(iii) the association of (18) with (10),

and (iv) the association of (18) with (11)

We proceed to discuss the modes of motion defined by each type

4—SOME APPROXIMATE RESULTS

The coasts of the strait are defined by the parameter a The more interesting cases occur when a is small compared with unity We therefore proceed to form approximations in which a is neglected in comparison with unity

Since in equation (6), $0 \leqslant \theta \leqslant 1$, for small a that equation becomes

$$\theta(1-\theta)\frac{d^2M}{d\theta^2} - \tfrac{1}{2}\theta\frac{dM}{d\theta} + \sigma(\sigma-1)\left(q-\frac{1}{\theta}\right)M = 0 \qquad (19)$$

The singularity $\theta = 0$ has indices σ, $1-\sigma$ Corresponding to the former the solution is

$$M = \Sigma a_n \theta^{\sigma+n},$$

where $a_{n+1}(n+2\sigma)(n+1) + a_n\{\sigma(\sigma-1)q - (\sigma+n)(\sigma+n-\frac{1}{2})\} = 0 \qquad (20)$

From this we find $\dfrac{a_{n+1}}{a_n} \sim 1 - \dfrac{3}{2n}$, and by the same argument as before it would lead to discontinuity of the ξ-component of velocity on the transverse axis

This series is only valid at $\theta = 1$ if it terminates, and this requires that

$$\sigma(\sigma-1)q = (\sigma+n)(\sigma+n-\tfrac{1}{2}), \qquad (21)$$

for some integer n

A further necessary condition is that $\sigma > 0$ Next consider equation (5 ii) Since $0 \leqslant \mu \leqslant a$, then provided that q is large compared with a, the equation reduces to

$$\mu(a-\mu)\frac{d^2N}{d\mu^2}+\tfrac{1}{2}(a-3\mu)\frac{dN}{d\mu}+\sigma(\sigma-1)qN = 0 \tag{22}$$

The indices at $\mu = 0$ are 0, $\frac{1}{2}$ Hence the solution is

$$N = \Sigma a'_n \mu^n,$$

where $$a'_{n+1}\; a(n+1)(n+\tfrac{1}{2})+a'_n\{\sigma(\sigma-1)q-n(n+\tfrac{1}{2})\} = 0, \tag{23}$$

or $$N = \Sigma a'_n \mu^{n+\frac{1}{2}},$$

where $$a'_{n+1}\; a(n+1)(n+\tfrac{3}{2})+a'_n\{\sigma(\sigma-1)q-(n+\tfrac{1}{2})(n+1)\} = 0 \tag{24}$$

Each of these series is valid up to $\mu = a$ only if it terminates That is

$$\sigma(\sigma-1)q = m(m+\tfrac{1}{2}), \quad \text{or} \quad (m+\tfrac{1}{2})(m+1), \tag{25}$$

where m is an integer

The solution (20) may be associated with (23) or with (24) to give approximately the modes of the first and second types

Lastly we may reduce (17) in the same way, obtaining

$$\theta(1-\theta)\frac{d^2w}{d\theta^2}-(1+\tfrac{1}{2}\theta)\frac{dw}{d\theta}+\left\{\frac{3}{2}\frac{1}{\theta}+\sigma(\sigma-1)\left(q-\frac{1}{\theta}\right)\right\}w = 0 \tag{26}$$

The requisite solution is

$$w = \Sigma a_n \theta^{\sigma+n+\frac{1}{2}},$$

where $$a_{n+1}(2\sigma+n)(n+1)+a_n\{\sigma(\sigma-1)q-(\sigma+n)(\sigma+n+\tfrac{1}{2})\} = 0, \tag{27}$$

and the corresponding solution of (5 i) is

$$M = \Sigma a_n \surd(1-\theta)\,\theta^{\sigma+n}$$

The solution determined by (27) is finite when

$$\sigma(\sigma-1)q = (\sigma+m)(\sigma+m+\tfrac{1}{2}), \tag{28}$$

m being an integer

By associating (27) with (23) and with (24) we may determine approximately the modes of the third and fourth types

We proceed to discuss the types in detail

(i) *First Type*

Let $p \equiv \sigma(\sigma-1)q$, and m, n be integers From (21) and (25) we have

$$p = m(m+\tfrac{1}{2}),$$

$$\sigma = m-n+\tfrac{1}{2}, \quad \text{or} \quad -m-n$$

Since σ must be positive the second value is discarded and $m \geqslant n$

Let $\mathrm{P}_n(x)$ be a polynomial in x of order n For this type the complete solution is

$$\text{Period} = \frac{2\pi}{\lambda} = \frac{\pi c}{\sqrt{(gh_0)}}\{m(m+\tfrac{1}{2})\}^{-\frac{1}{2}},$$

$$\zeta = e^{i\lambda t}(\cosh\xi)^{-2m+2n-1}\,\mathrm{P}_n(\cosh^{-2}\xi)\;\mathrm{P}'_m(\cos^2\eta), \tag{29}$$

P' being a second form of polynomial

It is readily shown in the usual way that $\mathrm{P}_n(x^2)$ has n real zeros in x The motion given by (29) therefore exhibits n nodal lines which are portions of ellipses and m nodal lines which are hyperbolas

It will appear later that when a is of the order 0 1 the approximation just found is fairly close to the real value except for the case $m = n = 0$

(ii) *Second Type*

From (21) and (25) we find

$$p = (m+\tfrac{1}{2})(m+1),$$

$$\sigma = m-n+1, \quad \text{or} \quad -m-n-\tfrac{1}{2}$$

The latter is rejected, and again $m \geqslant n$

The complete solution is now

$$\text{Period} = \frac{2\pi}{\lambda} = \frac{\pi c}{\sqrt{(gh_0)}}\{(m+\tfrac{1}{2})(m+1)\}^{\frac{1}{2}},$$

$$\zeta = e^{i\lambda t}(\cosh\xi)^{-2m+2n-2}\,\mathrm{P}_n(\cosh^{-2}\xi)\;\cos\eta\,\mathrm{P}'_m(\cos^2\eta)$$

This type of motion exhibits nodal lines similar to those of the first type, but in addition the conjugate axis is always a nodal line

(iii) *Third Type*

From (28) and (25) we find

$$p = m(m+\tfrac{1}{2}),$$

$$\sigma = m-n, \quad \text{or} \quad -m-n-\tfrac{1}{2}$$

The latter value is rejected and $m > n$

The complete solution for this type is

$$\text{Period} = \frac{2\pi}{\lambda} = \frac{\pi c}{\sqrt{(gh_0)}}\{m(m+\tfrac{1}{2})\}^{-\frac{1}{2}},$$

which is the same as the value for the first type to the degree of approximation used, and

$$\zeta = e^{i\lambda t}\sinh\xi(\cosh\xi)^{-2m+2n-1}\,\mathrm{P}_n(\cosh^{-2}\xi)\;\mathrm{P}'_m(\cos^2\eta)$$

The transverse axis is an additional nodal line

(iv) *Fourth Type*

$$p = (m+\tfrac{1}{2})(m+1),$$

$$\sigma = m-n+\tfrac{1}{2}, \quad m \geqslant n$$

$$\text{Period} = \frac{2\pi}{\lambda} = \frac{\pi c}{\sqrt{(gh_0)}}\{(m+\tfrac{1}{2})(m+1)\}^{-\frac{1}{2}},$$

which is the same as the value for the second type to the degree of approximation used, and

$$\zeta = e^{i\lambda t}\sinh\xi(\cosh\xi)^{-2m+2n-2}\,\mathrm{P}_n(\cosh^{-2}\xi)\;\cos\eta\,\mathrm{P}_m(\cos^2\eta)$$

The transverse axis and the conjugate axis are both additional nodal lines

It is noteworthy that in each of the solutions the presence of a factor in the form of a negative power of $\cosh\xi$ ensures that the wave amplitude diminishes with increasing ξ In the simplest cases the wave amplitude ultimately diminishes as $\cosh^{-1}\xi$ This result is important in connexion with the remarks already made on the possibility of applying the method of long waves

5—CALCULATION OF THE LOWER MODES OF VIBRATION

We proceed to calculate more exactly the lower modes of vibration of each type by continued approximation Only one numerical value need be assumed at the outset, namely a, the others, c and h_0 may be left in their general form

A convenient value for a is 0 1 The bounding hyperbola for the strait has thus an eccentricity 3 16, and the acute angle between the asymptotes is 36° 48′

The process is an extension of that of Hough (1897) The constants

p, σ are determined from the simultaneous equations (14) and (15), making a start by using the approximate results already found

(i) *First Type*

(*a*) *Lowest Mode corresponding to* $n = m = 0$

Assuming that $N_2 = a$, $N_2' = 1$, from (14) and (15) we have

$$\left.\begin{aligned} L_0 - \frac{y_1}{L_1 + a} &= 0, \\ L_0' - \frac{y_1'}{L_1' + 1} &= 0 \end{aligned}\right\} \tag{30}$$

Using the values $p = 0$, $\sigma = \frac{1}{2}$ in the second members and solving we have the next approximation

$$\left.\begin{aligned} p &= -0{\cdot}0096, \\ \sigma &= 0{\cdot}501 \end{aligned}\right\} \tag{31}$$

For the second approximation, take $N_4 = a$ and determine in succession the other N's with the values of p, σ in (31) We have then

$$N_3 = 0{\cdot}0739, \quad N_2 = 0{\cdot}0604, \quad N_1 = 0{\cdot}0334,$$

and finally
$$L_0 + N_1 = -0{\cdot}0019 \tag{32}$$

For correct values of p, σ, $L_0 + N_1$ should be zero In a similar manner, beginning with $N_4' = 1$, we find

$$N_3' = 0{\cdot}625, \quad N_2' = 0{\cdot}455, \quad N_1' = 0{\cdot}170,$$

and finally
$$L_0' + N_1' = -0{\cdot}021 \tag{33}$$

On solving (32) and (33) we have the improved values

$$p = -0{\cdot}0085, \quad \sigma = 0{\cdot}498$$

On repeating the operations with these values of p, σ we find as a closer approximation

$$\begin{aligned} p &= -0{\cdot}00863, \\ \sigma &= 0{\cdot}498 \end{aligned}$$

A final repetition gives the result, correct to three significant figures,

$$\begin{gathered} N_1 = 0{\cdot}0333, \quad N_2 = 0{\cdot}0601, \quad N_3 = 0{\cdot}0740, \\ N_1' = 0{\cdot}1719, \quad N_2' = 0{\cdot}4571, \quad N_3' = 0{\cdot}626, \\ p = -0{\cdot}00859, \quad \sigma = 0{\cdot}498 \end{gathered}$$

Hence we have the final results

$$\text{Period} = 7.81\ \pi c/\sqrt{(gh_0)},$$

$$\zeta = e^{i\lambda t}(\cosh\xi)^{-0.996}(1 + 0.0333\cosh^{-2}\xi + 0.0020\cosh^{-4}\xi + 0.0015\cosh^{-6}\xi \ldots)$$

$$\times(1 + 0.1719\cos^2\eta + 0.4571\cos^4\eta + 0.0492\cos^6\eta \ldots)$$

Since the maximum value of $\cos^2\eta$ is a, the second series converges rapidly.

It is noticeable that the expression for the displacement follows closely that indicated by the approximate solution (29). There are no nodal lines in the finite part of the sea. A nodal line is necessary and may be regarded as at infinity.

(*b*) *Second Mode corresponding to* $m = 1$, $n = 0$

$$\text{Period} = 0.848\ \pi c/\sqrt{(gh_0)}$$

$$\zeta = e^{i\lambda t}(\cosh\xi)^{-3\cdot 002}$$

$$\times(1 + 0.0859\cosh^{-2}\xi + 0.00718\cosh^{-4}\xi + 0.00061\cosh^{-6}\xi \ldots)$$

$$\times(-0.0340 + \cos^2\eta + 0.364\cos^4\eta + 0.214\cos^6\eta + 0.1472\cos^8\eta \ldots)$$

There are nodal lines given by the zero $\cos^2\eta = 0.0336$. These are the branches of the hyperbola of eccentricity 5.46.

The approximate solution (29) gives 0.816 as the numerical factor in the period. That solution also is limited to the first term of the $\cosh\xi$ series and the first two of the $\cos^2\eta$ series.

Figure 1 shows the wave-contours in a quadrant of the strait for this mode.

(*c*) *Third Mode corresponding to* $m = 1$, $n = 1$

$$\text{Period} = 0.834\ \pi c/\sqrt{(gh_0)}$$

$$\zeta = e^{i\lambda t}\cosh^{-1}\xi(-0.721 + \cosh^{-2}\xi + 0.0862\cosh^{-4}\xi + 0.00752\cosh^{-6}\xi \ldots)$$

$$\times(-0.0355 + \cos^2\eta + 0.654\cos^4\eta + 0.475\cos^6\eta + \ldots)$$

There are nodal lines given by $\cos^2\eta = 0.0347$ and by $\cosh^2\xi = 1.479$. The first gives the two branches of the hyperbola of eccentricity 5.37 and the second two portions of the ellipse of eccentricity 0.824.

(ii) *Second Type*

In place of (15) we take similar forms from (11).

Put, for this case,

$$L'_n = \frac{p-(n+\frac{1}{2})^2(1+a)-\frac{1}{2}(n+\frac{1}{2})}{a(n+1)(n+\frac{3}{2})},$$

$$y'_n = \frac{n^2-\frac{1}{4}-\sigma(\sigma-1)}{a(n+1)(n+\frac{3}{2})},$$

and consequently, as before,

$$\left.\begin{aligned} N'_{n+1}+L'_n+y'_n/N_n &= 0, \\ \text{and} \qquad N'_n &= -\frac{y'_n}{L'_n+N'_{n+1}} \end{aligned}\right\} \tag{34}$$

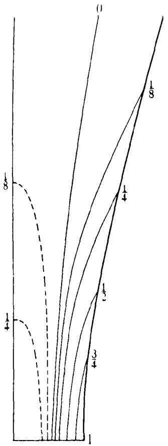


FIG 1—Showing the co range lines in a quadrant of the strait for the second mode of the first type

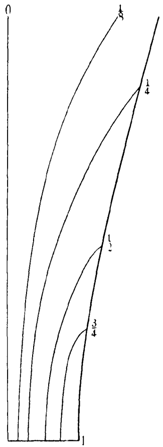


FIG 2—Showing the co range lines in a quadrant of the strait for the lowest mode of the second type

We now solve the pair (14) and (34) in the same manner as before

The lowest mode of this type is given by taking $m = n = 0$ We find then

$$\text{Period} = 1\ 424\ \ \pi c/\sqrt{(gh_0)}$$

$$\zeta = e^{i\lambda t}(\cosh \xi)^{-1\ 978}$$

$$\times (1 + 0\ 0586 \cosh^{-2} \xi + 0\ 00415 \cosh^{-4} \xi + 0\ 00033 \cosh^{-6} \xi \quad)$$

$$\times \cos \eta (1 + 0\ 307 \cos^2 \eta + 0\ 166 \cos^4 \eta + 0\ 113 \cos^6 \eta \quad)$$

This mode is distinguished by having a nodal line along the conjugate axis In fact, the motion is asymmetrical with regard to this line

Figure 2 shows the wave contours in a quadrant of the strait for this mode

(iii) *Third Type*

The recurrence formula (18) may similarly be used with (15) for the third type of motion

The lowest mode of this type is found by taking $m = 1$, $n = 0$ We have then

$$\text{Period} = 0\ 837\ \ \pi c/\sqrt{(gh_0)}$$

$$\zeta = e^{i\lambda t} \sinh \xi (\cosh \xi)^{\ 3\ 002}$$

$$\times (1 + 0\ 0865 \cosh^{-2} \xi + 0\ 00727 \cosh^{-4} \xi + 0\ 00063 \cosh^{-6} \xi + \quad)$$

$$\times (-0\ 03497 + \cos^2 \eta + 0\ 5804 \cos^4 \eta + 0\ 4063 \cos^6 \eta + \quad)$$

The transverse axis is a nodal line and the motion is asymmetrical with regard to it There is also a pair of nodal lines given by the hyperbola $\cos^2 \eta = 0\ 0343$

Figure 3 shows the wave contours in a quadrant of the strait for this mode

(iv) *Fourth Type*

For this type we use (18) and (11)

The lowest mode is given by $m = n = 0$

$$\text{Period} = 1\ 428\ \ \pi c/\sqrt{(gh_0)}$$

$$\zeta = e^{i\lambda t} \sinh \xi (\cosh \xi)^{-2\ 108}$$

$$\times (1 + 0\ 0611 \cosh^{-2} \xi + 0\ 00443 \cosh^{-4} \xi + 0\ 00035 \cosh^{-6} \xi \quad)$$

$$\times \cos \eta (1 + 0\ 403 \cos^2 \eta + 0\ 232 \cos^4 \eta + 0\ 161 \cos^6 \eta + \quad)$$

Both the transverse axis and the conjugate axis are nodal lines

The contours for this mode are shown in Figure 4

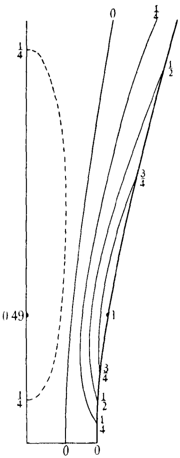


FIG 3—Showing the co range lines in a quadrant of the strait for the lowest mode of the third type

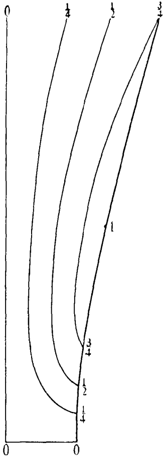


FIG 4—Showing the co range lines in a quadrant of the strait for the lowest mode of the fourth type

I am indebted to Mr John L Scott for assistance in preparing the diagrams

6—SUMMARY

The problem studied is that of the natural vibrations of the water in a strait connecting two open seas The shores of the strait are the two branches of a hyperbola and a certain law of depth is chosen It is shown that such a system has free modes of four distinct types characterised by symmetry about both axes, asymmetry about the conjugate axis, asymmetry about the transverse axis and asymmetry about both axes, respectively In each case there is an infinite number of distinct modes The waves diminish rapidly in amplitude as the channel widens The nodal lines are members of the same family of confocal ellipses and hyperbolas of which the shores are members

Some of the lower modes are worked out completely for special cases But simpler approximations are given covering all the forms where the eccentricity of the bounding hyperbola is not too small

REFERENCES

Goldsbrough, G R 1930 *Proc Roy Soc* A, **130** 157–67

— 1936 *Proc Roy Soc* A, **155**, 12 32

Harris, R A 1908 "Manual of Tides', Part v Washington Government Printing Office

Honda, Terada, Yoshida and Isitani 1908 *J Coll Sci Tokyo*, **24** 1–113

Hough, S S 1897 *Phil Trans* A **189**, 201

Lamb 1932 "Hydrodynamics', 6th Ed p 344

Production and Dissipation of Vorticity in a Turbulent Fluid

By G I Taylor, F R S

(*Received* 6 *October* 1937)

In some recent publications (Taylor 1937, Taylor and Green 1937) the author has put forward the view that the high average vorticity which is known to exist in turbulent motion is caused by the extension of vortex filaments in an eddying fluid Let A and B be two particles a short distance, d_0, apart on a vortex line where the vorticity is ω_0 At a subsequent time when the distance between A and B has changed from d_0 to d and the vorticity from ω_0 to ω then, neglecting the effects of viscosity, the equation representing the conservation of circulation is

$$\omega/\omega_0 = d/d_0 \tag{1}$$

Turbulent motion is found to be diffusive, so that particles which were originally neighbours move apart as the motion proceeds In a diffusive motion the average value of d^2/d_0^2 continually increases It will be seen therefore from (1) that the average value of ω^2/ω_0^2 continually increases An equation for the average rate of increase in ω^2 has been given by v Karman (1937) which contains the term $\overline{\omega^2 \frac{\partial u_3}{\partial x_3}}$, where the bar shows that the average value has been taken and $\partial u_3/\partial x_3$ represents the rate of stretching of vortex filaments

The argument given above would lead to the expectation that there is a positive correlation between ω^2 and $\partial u_3/\partial x_3$, since places where the vortex filaments are stretching are places where high vorticities may be expected On the other hand, v Karman (1937) made the contrary assumption that $\overline{\omega^2 \frac{\partial u_3}{\partial x_3}} = 0$

In conversation with the author v Karman pointed out that apart from the quantity under discussion, namely, $\overline{\omega^2 \frac{\partial u_3}{\partial x_3}}$, his equation contains only quantities which can be measured in a wind tunnel, so that it should be possible to estimate the value of $\overline{\omega^2 \frac{\partial u_3}{\partial x_3}}$ in an actual case of turbulent flow This suggestion has now been carried into effect

v Karman's equation is

$$\frac{\partial}{\partial t}\overline{\omega^2} = 2\nu \Sigma_i \Sigma_k \overline{\left(\frac{\partial \omega_i}{\partial x_k}\right)^2} + 2S, \tag{2}$$

where x_i, x_j, x_k are the co-ordinates and ω_i, ω_j, ω_k are the components of vorticity so that $\omega_i^2 + \omega_j^2 + \omega_k^2 = \omega^2$, S is the quantity under discussion, namely (v Karman 1937),

$$S = \overline{\omega^2 \frac{\partial u_3}{\partial x_3}} = \Sigma_i \Sigma_k \overline{\omega_i \omega_k \frac{\partial u_i}{\partial x_k}} \tag{3}$$

The components of (2) will next be expressed in terms of quantities which have been measured in a wind tunnel These measurements were

1—The values of $\overline{u^2}$ at various distances x down-stream from a 3×3 in square-mesh grid

2—The correlation, R_1, between the velocities at points r apart in a line parallel to the direction of the wind

These measurements have already been published (Taylor 1937, figs 2 and 3) Since the average conditions at a fixed point in the tunnel are constant while the whole turbulence is carried by a stream of mean velocity U, $\frac{\partial}{\partial t}$ in (2) corresponds with $U\frac{\partial}{\partial x}$ applied to the measured quantities

Consider the first term in (2), namely, $\frac{\partial}{\partial t}\omega^2$ It has been shown that (Taylor 1937) in these measurements the turbulence was isotropic, and for isotropic turbulence (Taylor 1935)

$$\overline{\omega^2} = 15\overline{u^2}/\lambda^2, \tag{4}$$

where

$$\lambda^{\;2} = \left[\frac{\partial^2 R_1}{\partial r^2}\right]_{r=0} \tag{5}$$

λ is related to the variation in $\overline{u^2}$ with x thus (Taylor 1935 and v Karman 1937)

$$U\frac{\partial}{\partial x}(\overline{u^2}) = \frac{\partial}{\partial t}(\overline{u^2}) = -10\frac{\nu\overline{u^2}}{\lambda^2}, \tag{6}$$

where $\nu = \mu/\rho$ is the dynamic viscosity Combining (4) and (5)

$$\frac{\partial}{\partial t}\overline{\omega^2} = 15\frac{\partial}{\partial t}\left(\frac{\overline{u^2}}{\lambda^2}\right) = -\frac{3}{2\nu}\frac{\partial^2}{\partial t^2}(\overline{u^2}) \tag{7}$$

In representing the measured values of $\overline{u^2}$ on a diagram it is convenient to plot $U/\sqrt{\overline{u^2}}$ against x because it has been found in some cases (and in particular for the present measurements) that the curve so produced is a straight line (see Taylor 1937, fig 3) If β is the slope of the $(U/\sqrt{\overline{u^2}}, x)$ curve, β may be regarded as a quantity which has been measured for a range of values of x In the present measurements β, as mentioned above, was independent of x, but the analysis applies also in cases when β is a function of x

Writing

$$\beta = \frac{d}{dx}\left(\frac{U}{u'}\right) = \frac{\partial}{\partial t}\left(\frac{1}{u'}\right), \tag{8}$$

where

$$u' = \sqrt{\overline{u^2}}$$

it will be noticed that

$$\left.\begin{aligned} \frac{\partial}{\partial t}(\overline{u^2}) &= -2u'^3\beta, \\ \text{and}\quad \frac{\partial^2}{\partial t^2}(\overline{u^2}) &= 6\beta^2u'^4 - 2Uu'^3\frac{d\beta}{dx} \end{aligned}\right\} \tag{9}$$

Hence from (7)

$$\frac{\partial}{\partial t}(\overline{\omega^2}) = -9\frac{u'^4\beta^2}{\nu} + \frac{3u'^3U}{\nu}\frac{d\beta}{dx} \tag{10}$$

The first term on the right-hand side of (2) is associated with $\left[\frac{d^4R_1}{dr^4}\right]_{r=0}$ by an equation which is analogous to that which connects $\overline{\left(\frac{\partial u}{\partial x}\right)^2}$ with $\left[\frac{d^2R_1}{dx^2}\right]_{r=0}$, namely (Karman 1937),

$$\Sigma_i\Sigma_k\overline{\left(\frac{\partial\omega_i}{\partial x_k}\right)^2} = \frac{7}{3}\overline{\omega^2}\lambda^2\left[\frac{d^4R_1}{dr^4}\right]_{r=0} \tag{11}$$

Substitution from (3), (10) and (11) in (2)

$$S = -\frac{9u'^4}{2\nu}\beta^2 + \frac{3u'^3}{\nu}U\frac{d\beta}{dx} + 35\nu\overline{u^2}\left[\frac{d^4R_1}{dr^4}\right]_{r=0} \tag{12}$$

EXPERIMENTAL DATA

The measurements here described were made by Mr L F G Simmons at the National Physical Laboratory in a wind tunnel across which a 3×3 in square mesh grid was placed It was found that when U/u' was plotted against x a good straight line was attained so that $\frac{d\beta}{dx} = 0$

The value of U/u' at $x = 6$ ft 10 in was 33 and at 14 ft 10 in it was 69, so that $\beta = \frac{36}{96\times2\cdot54} = 0\cdot147\text{ cm}^{-1}$ (Taylor 1937, p 314) At $U = 15$ ft /sec, and at $x = 6$ ft 10 in therefore $u' = \frac{15}{33}\times12\times2\cdot54 = 13\cdot8$ cm /sec Using $\nu = 0\cdot14$ c g s it will be found that

$$-\frac{1}{2}\frac{\partial}{\partial t}\overline{\omega^2} = \frac{9}{2}\frac{u'^4}{\nu}\beta^2 = 2\cdot50\times10^4\text{ sec}^{-3} \tag{13}$$

Evaluation of $\left[\frac{d^4R_1}{dr^4}\right]_{r=0}$

The observed values at $U = 15$ ft /sec of R_1 near $r = 0$ were plotted on a large scale, and the following values were taken from the faired curve (fig 1) drawn as nearly as possible through the observed points

To find $\left[\frac{d^4R_1}{dr^4}\right]_{r=0}$ one may determine three constants a, b and c in the expression

$$1 - R_1 = ar^2 + br^4 + cr^6, \tag{14}$$

so that it fits three of the observations given in Table I Then evidently

$$\lambda^{-2} = -\left[\frac{d^2 R_1}{dr^2}\right]_{r=0} = 2a, \tag{15}$$

and

$$\left[\frac{d^4 R_1}{dr^4}\right]_{r=0} = -24b \tag{16}$$

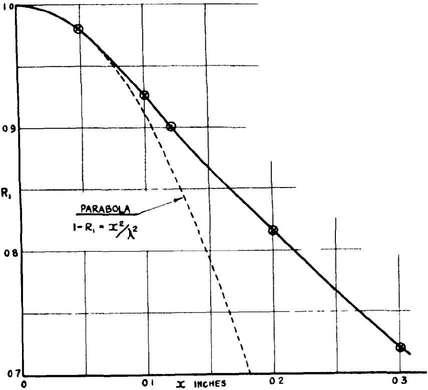


Fig 1—Faired curve for observed values of R_1

TABLE I

Inches	$1-R_1$
0 10	0 075
0 12	0 100
0 15	0 133
0 20	0 185
0 25	0 233
0 30	0 282

Inserting in (14) the values of $(1-R_1)$ given in Table I for $r = 0{\cdot}1$, $0{\cdot}15$ and $0{\cdot}2$ in and solving for a, b and c the following values are obtained

$$\left.\begin{aligned} a &= 9\ 15\,\text{in}^{-2}, \\ b &= -182\,\text{in}^{-4}, \\ c &= 1730\,\text{in}^{-6} \end{aligned}\right\} \tag{17}$$

Inserting this value of a in (15)

$$\lambda = (2a)^{-\frac{1}{2}} = 0\ 234\,\text{in} = 0{\cdot}59\,\text{cm} \tag{18}$$

This value may be checked against the value derived from the equation for the decay of turbulence which may be written in the form

$$\lambda = \sqrt{\frac{5\nu}{\beta u'}} \tag{19}$$

Inserting $\nu = 0\ 14$, $\beta = 0\ 147$, $u' = 13\ 8$ cm /sec , it is found that

$$\lambda = 0\ 59\,\text{cm} \tag{20}$$

The agreement between (18) and (20) is probably better than is warranted by the experimental data

The parabola which coincides with the R_1 curve at its vertex is shown by a broken line in fig 1 It is clear that the method described above for determining $\left[\dfrac{d^2R_1}{dr^2}\right]_{r=0}$ and $\left[\dfrac{d^4R_1}{dr^4}\right]_{r=0}$ is only accurate if the range of the R_1 curve used is so small that it can be represented by a small number of terms of the expansion in even powers of r It is worth while setting down the values of the separate terms in the series (14) for some values of r intermediate between the 3 values = 0 1, 0 15 and 0·2 in which were used to determine a, b and c This is done in Table II

TABLE II—$a = 9\ 15$, $b = -182$, $c = 1730$

Inches	ar^2	br^4	cr^6	$ar^2+br^4+cr^6$	$1-R_1$ (fig 1)
0 05	0 0229	−0 0011	0 00003	0 0218	0 022
0 10	0 0915	−0 0182	0 00173	0 0750	0 075
0 13	0 1543	−0 0519	0 00833	0 1107	0 111
0 15	0 2060	−0 0923	0 01970	0 133	0 133
0 17	0 2645	−0 1520	0 04150	0 154	0 154
0 20	0 366	−0 2900	0 11000	0 186	0 186
0 25	0 572	−0 7100	0 425	0 287	0 233

In cols 2, 3 and 4 are the values of ar^2, br^4, cr^6, while in col 5 the values of $ar^2+br^4+cr^6$ are given In col 6 are shown the values of R_1 taken from fig 1 The figures in cols 5 and 6 are identical at $r = 0·1$, $0·15$ and $0·2$ in because the constants a, b and c are determined from these three points, but the coincidence at $r = 0·05$, $0·15$, $0·17$ shows that the approximation to r^6 is sufficiently accurate to cover the range from $r = 0$ to $r = 0·2$ There is considerable difference between cols 5 and 6 at $r = 0·25$, due no doubt to the fact that terms in higher powers of r are becoming appreciable

Substituting from (17) in (16)

$$\left[\frac{d^4R_1}{dr^4}\right]_{r=0} = 24 \times 182 \text{ in}^{-4} = 105 \text{ cm}^{-4}, \tag{21}$$

so that

$$35\nu\bar{u^2}\left[\frac{d^4R_1}{dr^4}\right]_{r=0} = 35 \times 0·14 \times (1·38)^2 \times 1050 = 9·8 \times 10^4 \tag{22}$$

Substituting from (22) and (13) in (12) it will be seen that in the present case

$$S = -2·50 \times 10^4 + 9·8 \times 10^4 = 7·3 \times 10^4 \text{ sec}^{-3} \tag{23}$$

The first term, $2·5 \times 10^4$, represents the rate at which vorticity is decreasing the second term, $9·8 \times 10^4$, represents the rate at which vorticity is being destroyed by viscosity It will be seen that in this case the rate at which vorticity is being destroyed by viscosity is four times as great as the rate at which it is disappearing in the turbulent field The deficiency, equal to three times the rate of decay of vorticity, is supplied by the action represented by S namely the extension of vortex filaments in places where the vorticity is high

Effect of Increase in Wind Speed

As the speed U of the flow increases $\left[\frac{d^2R_1}{dr^2}\right]_{r=0}$ increases proportionally to U, and if U is sufficiently great the (R_1, r) curve becomes nearly pointed at the top Even in the limit when $U \to \infty$ this part of the curve however seems to be a finite angle rather than a cusp Measuring the slope of the approximately straight part of the curve in fig 1, i e the portion between $r = 0·1$ in and $r = 0·3$ in, it is found that the angle in this case corresponds with a value of $\frac{dR_1}{dr}$ equal to $-1·02 \text{ in}^{-1}$

The portion of the R_1 curve shown in fig 1 may be regarded roughly as being made up of a parabolic part up to $r = 0·06$ in and a straight line from $r = 0·06$ to $r = 0·2$ in

It will now be assumed that as U increases the upper part of the R_1 curve consists of a parabola joining on smoothly to a straight line whose slope is

$$\frac{dR_1}{dr} = 1\ 02\,\text{in}^{-1} \tag{24}$$

At the point where the straight line is tangential to the parabola, namely, $r = 0\ 06$ in the term br^4 in (14) is becoming appreciable compared with ar^2, in fact at $r = 0\ 06$

$$\frac{br^4}{ar^2} = 0\ 07 \tag{25}$$

We can now derive roughly the value of $\left[\frac{d^4R_1}{dr^4}\right]_{r-0}$ in terms of $\left[\frac{d^2R_1}{dr^2}\right]_{r=0}$ as follows

If r_1 is the value of r at which the straight line is tangential to the parabola we may assume that $\frac{br_1^4}{ar_1^2} = 0{\cdot}07$

so that
$$\left[\frac{d^4R_1}{dr^4}\right]_{r-0} = \frac{12(0{\cdot}07)}{r_1^2}\left[\frac{d^2R_1}{dr^2}\right]_{r-0} \tag{26}$$

To find r_1, notice as in (24) that the shape of the parabola at $r - r_1$ must be such that $\frac{d}{dr}(R_1) = -1\ 02$, thus $r_1\left[\frac{d^2R_1}{dr^2}\right]_{r=0} = -1\ 02$

It appears therefore that in this case

$$\left[\frac{d^4R_1}{dr^4}\right]_{r-0} = \frac{12(0\ 07)}{(1\ 02)^2}\left\{\left[\frac{d^2R_1}{dr^2}\right]_{r=0}\right\}^3 \tag{27}$$

Taking, for instance, the case where $U = 15$ ft /sec

$$\left[\frac{d^2R_1}{dr^2}\right]_{r-0} = 2 \times 9\ 15 = 18\ 3\,\text{in}^{-2},$$

so that the approximate formula (24) gives

$$\left[\frac{d^4R_1}{dr^4}\right]_{r-0} = 0\ 81(18\ 3)^3 = \underline{4\ 9 \times 10}^3\,\text{in}^{-4}$$

The true value given in (21) is $24 \times 182 = \underline{4\ 4} \times 1\underline{0}^3\,\text{in}^{-4}$ Thus the value obtained by using the rough approximation is not far from the true value

Using this approximation the effect of change in wind speed can be estimated Since the value of u'/U at a fixed point remains constant as U

varies, it will be seen from (22) that $\lambda^{-2}\propto U$ and since $\lambda^{-2} = \left[\frac{d^2R_1}{dr^2}\right]_{r=0}$ (27) shows that

$$\left[\frac{d^4R_1}{dr^4}\right]_{r=0} \propto U^3 \qquad (28)$$

Sufficient data are now available to predict the effect of an increase in U on the terms in Karman's equation (12) for rate of change in vorticity Referring to (12) and remembering that u'/U was found to be nearly constant as U changed, it will be seen that the first term $\frac{1}{2}\frac{\partial}{\partial t}\overline{\omega^2}$ is proportional to U^4, while the second term is proportional to $U^2\left[\frac{d^4R_1}{dr^4}\right]_{r=0}$, i e to U^5

It appears therefore that the ratio of the rate of dissipation of vorticity by viscosity to the rate of decrease in vorticity increases proportionally to U When U is very great rate of production of vorticity by the extension of vortex filaments becomes very great compared with the rate at which vorticity is decreasing

SUMMARY

When isotropic turbulence is set up in a fluid, e g by moving a grid of regularly spaced bars through it the average vorticity decreases with time, this decrease is due to viscosity Recently von Karman has calculated the rate at which vorticity is destroyed by viscosity His equation involves only quantities which can be measured in a wind tunnel by means of the hot wire technique These quantities have now been measured in one case, and the rate of decrease in the mean square of the vorticity has also been measured In this case it is proved that the rate of destruction of vorticity by viscosity is four times as great as the rate at which vorticity disappears Vorticity is therefore being produced by extension of vortex filaments three times as fast as it is disappearing

It seems that the stretching of vortex filaments must be regarded as the principal mechanical cause of the high rate of dissipation which is associated with turbulent motion

REFERENCES

Karman, O v 1937 *J Aeronaut Sci* **4**
Taylor, G I 1935 *Proc Roy Soc* A **151** 421
— 1937 *J Aeronaut Sci* **4**, 311
Taylor, G I and Green, A E 1937 *Proc Roy Soc* A, **158**, 499

Galvano-magnetic Effects in Bismuth Alloys

By N Thompson, *H H Wills Physical Laboratory, University of Bristol*

(*Communicated by A M Tyndall, F R S —Received* 3 *July* 1937)

1—Introduction

There are few phenomena of comparable importance to which so much attention has been devoted as to galvano magnetic effects in bismuth, and few also where the results are so discordant A bibliography of the earlier work can be found in Campbell's book on the subject (1923), but in the light of present knowledge most of it is seen to be of little value, since the experiments were performed with polycrystalline material—and bismuth is above all things anisotropic References are given at the end of this article to those papers which deal with single crystals, but even these are not much use, since the metal used was, by modern standards, anything but pure, and the impurities not known Exceptions to this are the more recent papers by Kapitza (1928), Schubnikov and de Haas (1930), and de Haas, Gerritsen and Capel (1936), but each one of these covers only a limited range either of temperature or of magnetic field

In short, it is seen that no complete set of data exists for pure bismuth over an extended range of fields and temperatures, and that the effects of impurities present in the bismuth are only known in a vaguely qualitative way, for example, the various workers give values for the magneto-resistance coefficient of bismuth which, for the same values of field strength and temperature, differ by as much as a factor of twenty times

Some recent measurements by the author (1936) on the resistivity of impure bismuth suggested that the presence of small quantities of tetravalent or hexavalent impurity might be sufficient to account for these divergencies The question has been considered theoretically by Jones (1936), and it was mainly with the object of checking the findings of this theory that the present systematic investigation of the effects of traces of lead and selenium on the magneto-resistance effect and the Hall effect in bismuth was begun

2—Experimental

The specimens were prepared in much the same manner as described in the previous paper (Thompson 1936), the base metal used being again Hilger bismuth (No 10,283) with a quoted impurity content of Ag 0·001 % Pb 0·0004 % The device for casting the alloys into a suitable shape was modified slightly to enable the process to be carried out *in vacuo* The castings measured about 20 × 4 × 0·8 mm and were grown into single crystals of predetermined orientation as before The surface tension of the molten metal was sufficient to spoil the laminar shape of the specimen necessary for Hall effect measurements unless steps were taken to prevent it doing so This could be done by resting it in a shallow steatite mould, slightly larger than itself, on the top of which another thin slip of steatite formed a lid This served to keep the two faces parallel, without being a sufficient constraint to prevent the growth of a single crystal Next, for a distance of about 2 mm at either end, the crystal was coated electrolytically with copper, and on to this the two current leads were soldered For potential leads, 44 s w g copper wires were again used Two such, applied on the centre line of one of the faces, a few millimetres from the copper plating, served to measure the resistance, while two more, opposite one another about the centre of the long edges, gave the Hall e m f The crystal was supported vertically by the current leads inside a Dewar vessel between the poles of a powerful electromagnet It could be rotated about a vertical axis, and the amount of the rotation read on a scale at the top The magnetic field was found to be homogeneous to 0·1 % over a region considerably larger than that occupied by the specimen Fields up to 22,000 oersteds were obtainable

3—Examination of possible errors

In making observations of this kind, care must be taken to ensure that the temperature is uniform over the specimen (to avoid parasitic thermo-e m f 's), and that it is strictly constant during one set of observations The specimen must be mounted so that it cannot move under the action of the magnetic forces, and the current through it must not be so big as to heat it above the temperature of its surroundings Then, in addition to the ordinary errors which may be present in measuring the potentials, field strengths, temperatures, and specimen thicknesses observed, we must consider the following possibilities

a—Small temperature gradients may still persist along the specimen, and errors due to this cause can be eliminated by reversing the current through the specimen and taking the mean of the two readings

b—The reversible (Thomson) heating at the points where the current enters and leaves the specimen may alter the temperature at the points of contact of the potential leads, and so give rise to a differential thermo-e m f This will not be eliminated by reversing the current, but can be detected as an apparent change in resistance on suddenly doing so, followed by a drift back to the original value An effect corresponding to this description was occasionally observed if the potential leads were placed too near to the ends of the specimen

c—Galvano-magnetic temperature differences may be set up both along and across the specimen (e g the Ettinghausen effect), and these acting on the copper potential leads as differential thermocouples will give rise to parasitic e m f 's These e m f 's were expected to be small, and were shown to be negligible by the fact that when the copper potential leads were on one occasion replaced by constantan, the results were unaffected

d—Temperature inequalities in the specimen, however caused, will give rise direct to thermo-magnetic e m f 's (e g Nernst effect) Since both the temperature gradients likely to exist, and the parameters in the appropriate equations are small, this possibility can be safely ignored

e—In zero field, the Hall electrodes may not lie exactly on an equipotential line, and the resistance electrodes not on a line of current flow Consequently, when the field is switched on, there will be a component of the Hall e m f appearing at the resistance electrodes and vice versa This error is far more important than those previously mentioned, since it is only in exceptional cases that it is absent, and may be quite large Fortunately it can be eliminated merely by reversing the magnetic field (and interchanging the leads from the Hall electrodes) and taking the mean This procedure was invariably followed in practice, and one important consequence must be noted It follows immediately that the Hall e m f as so measured must be an odd function of the magnetic field, and the magneto-resistance effect an even function If the real effects are not so expressible (which appears unlikely, at least to a good first approximation) then only those components which are so expressible have been measured

f—It is well known that all galvano-magnetic phenomena in bismuth show a considerable variation with the relative orientation of the three vectors I (current), H (magnetic field) and oz (principal axis of the crystal) In all the measurements here described I and H are at right angles ("trans verse" effects) and oz is either parallel to one or the other of them, or else perpendicular to both These are the cases considered theoretically by Jones (1936), and, as Stierstadt (1933) has shown, represent in general (but see (*g*)) either a maximum or a minimum position This leads us to the next source of error, namely, that the crystal axis may not be correctly orientated This may happen either (1) because the crystal is not correctly grown, i e the angle between I and oz is wrong, or (2) because it is inaccurately placed in the field, i e the angle between H and oz is wrong The former point was tested by observing the cleavage plane perpendicular to oz and could be relied upon to one or two degrees The latter was done by adjusting the specimen to be vertical rotating the mounting about a vertical axis, and setting on a maximum (or minimum where appropriate) position for the magneto-resistance effect, usually at 90° abs and 14,000 oersteds This again could be done to one or two degrees Any serious error in crystal orientation, caused say by twinning would have made itself very obvious by the shape of the curve showing the variation of $\Delta\rho/\rho$ during this rotation, where ρ is the resistance in zero field at temperature T° abs and $\Delta\rho$ the change of resistance caused by switching on the magnetic field The error caused by a small misorientation of the axis (θ) is not easily calculable, under all conditions, but it is unlikely that the variation will be as rapid as $\cos\theta$, and even this means only a 1 % error for a misorientation as improbably large as 8°

q—The experiments of Stierstadt (1933) and of Gruneisen and Gielessen (1936) have shown that, for a given orientation of the principal axis, considerable variations can be produced by a rotation around that axis, that is, the orientation of the binary axes is of importance As yet no theory explains or even takes cognisance of this fact so that the best that could be done was to ensure that all the measurements made were consistent among themselves in this respect Consider first a rectangular plate of bismuth in the form of a single crystal with its principal axis (oz) normal to its surface and let it carry a current along its length which is to be kept always perpendicular to the magnetic field Then its resistance depends on the angle between H and oz ($= \phi$ say) The variation is symmetrical about a maximum (or minimum) value at $\phi = 0$ if and only if, one of the binary axes (oy) lies parallel to the current flow By careful manipulation, and the growing of one seed crystal from another, it was possible to produce the specimens with a binary axis not more than one or two degrees from the position just described The "rotation curves" of $\Delta\rho/\rho$ vs ϕ obtained with these were never far from perfect symmetry, and were in fact used as the method of orientating the specimen in the magnetic field The position $\phi = 0$ is referred to in Jones' paper as case "*d*" Similar considerations of symmetry for crystals whose principal axis lay along their length (Jones' case "*c*") led to a choice of the position in which a binary axis was normal to the plane of the lamina, and all the measurements were made on crystals so orientated Thus when set up for observations on the Hall effect, in the first case we have oz parallel to H and oy parallel to I in the second, oz parallel to I and oy parallel to H, the Hall e m f being in both instances observed in a direction normal to the plane containing H, I, oz and oy A rotation of 90° about a vertical axis from case "d" gives case "e" (oz perpendicular to both H and I, oy parallel to I) Only resistance measurements were made in this position

h—Lastly, slight errors in the orientation of the specimen in the field, or in the location of the potential leads on the specimen, may lead to a component of some "longitudinal' effect being measured along with the "transverse' effects with which we are concerned Since the former are usually of a lesser order of magnitude than the latter, and since, *a priori*, only a small resolved part could be present as an error, the errors introduced by this factor will only be small

In spite of this rather forbidding and still incomplete array of errors, it is likely that they do not add up to more than about 5 % in any one measurement For very small fields, when the relevant e m f's (or changes in e m f in the case of the magneto-resistance effect) are $10\,\mu$V or less, the causes a and b make the accuracy less, and extrapolations to zero field where these are done, are consequently less reliable still And, as previously, one of the biggest uncertainties lies in the fixing of the exact amount of impurity present in the bismuth The above list will at least serve to show why, at this stage, no attempt has been made at precision measurements

4—RESULTS ON MAGNETO-RESISTANCE EFFECT

In the past, experimental results on the magneto resistance effect have been displayed by the drawing of a graph showing the relative change of resistance as a function of the magnetic field When this is done, it appears that the curve is at first parabolic, and for higher fields linear However, in the paper by Jones referred to above, an expression is deduced for $\Delta\rho/\rho$ as a function of H which, while having the above form at the outset, gives curves which eventually become concave to the axis of H, and Jones shows that Kapitza's results for very high fields do fit an equation of this form Jones' equation, reproduced here for convenience of reference, is

$$\frac{\Delta\rho}{\rho}=\frac{BH^2}{1+z^2CH^2} \tag{1}$$

Here, z is the atomic percent of impurity present in the bismuth, and is to be considered positive or negative according as the impurity is 4- or 6-valent B and C are constants independent of H but involving z as a second-order effect, they vary rapidly with temperature, becoming larger as the temperature falls As can easily be seen, when $z^2CH^2 \ll 1$ (1) gives a parabola, while for $z^2CH^2 \gg 1$ it has an asymptote parallel to the axis of H The concavity shown by Kapitza's curves at high fields is the beginning of the approach to this asymptote It thus appears that the important quantity is not merely H itself, but the combination z^2CH^2, and that Kapitza's results in high fields should be reproducible in lower fields, either by increasing z, the impurity content, or by increasing C, i e by lowering the temperature The first result of this investigation is to show that qualitatively this is correct, as can be seen by comparing our curves with those of Kapitza Fig 1 shows

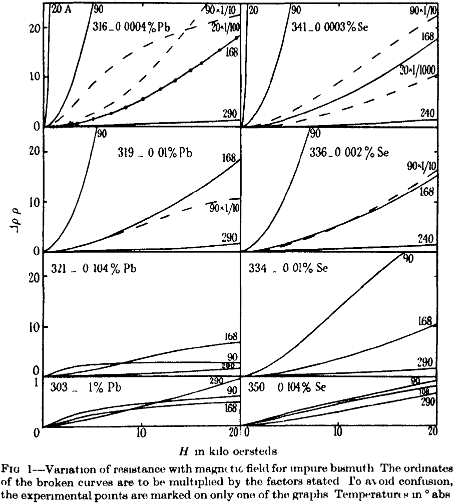


FIG 1—Variation of resistance with magnetic field for impure bismuth The ordinates of the broken curves are to be multiplied by the factors stated To avoid confusion, the experimental points are marked on only one of the graphs Temperatures in ° abs Impurity content in atomic percent

a selection of all the curves obtained, chosen to give a general idea of the run of the phenomena *

To obtain a more careful comparison between theory and experiment, it is convenient to rewrite equation (1) in the form

$$\frac{H^2\rho}{\Delta\rho} = \frac{1}{B} + \frac{z^2C}{B} H^2 \qquad (2)$$

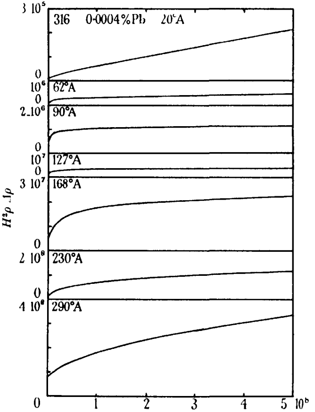


FIG 2—Bi + 0 0004 % Pb at 20–290° abs

Writing y for $H^2\rho/\Delta\rho$, and x for H^2, we see that we should, according to the theory, obtain a straight-line graph Some representative curves actually obtained are shown in fig 2 In all, some 150 such curves have been obtained,

---

* In examining this and the other figures in this paper, close attention should be paid to the vertical scales, as these often vary considerably as between one group and the next

at temperatures from 14 to 350° abs, but mainly at 65, 90, 125, 168, 230 and 290° abs, and for a range of alloys of bismuth with small amounts of lead and selenium It is obviously impracticable to reproduce all the results here, but it can be said that they all support the following general conclusions

1—For z very small, the experimental points lie on a straight line when both H and C are large (high fields and low temperatures)

2—As the impurity content is increased this ceases to be true for the fields and temperatures obtainable, while the indications are that it still holds for higher fields

3—The deviations from the straight line which are always present at low fields are always of the form shown, i e $H^2\rho/\Delta\rho$ lies below the straight line determined by the high-field values

4—The shape of the curves is the same whether lead or selenium is the impurity present, i e the phenomena depend on z^2 rather than on z

5—The above conclusions are true for all the crystal orientations dealt with

Thus on this second and closer examination it is seen that Jones's theory is true only for certain limiting cases Moreover, since the deviations from it occur for z large, H small and C small (T large), i e depend not on the combination z^2CH^2 which we have previously met, but rather on CH^2/z^2 some explanation of them must be sought other than merely refining the existing theory

At this point it might perhaps be opportune to add a note on fig 2, 90° abs curve This shows that the quantity y is still varying rapidly in the region of small x As previously mentioned, the experimental accuracy is much reduced in this region, and so the extrapolation to zero field was not very trustworthy, and indeed it was not certain that it was justified at all To settle this point a second crystal also of pure bismuth was grown this time in the form of a rod about 1 mm section and 10 cm long This was useless for Hall effect observations, but enabled the resistance changes in low fields to be measured with much greater accuracy The electromagnet was not suited to the production and measurement of small fields, and accordingly a Helmholtz double coil was built capable of giving fields up to 230 oersteds homogeneous to 0·1 % over a volume 10 cm diameter and 5 cm thick Using this arrangement, the variation of resistance with field was investigated at 90° abs for a crystal whose principal axis was parallel to its length After the measurements had been made, the specimen was carefully transferred to the electromagnet, and observations taken in higher fields The two sets of results are shown together in fig 3 It is apparent that there is no anomalous behaviour in very low fields, and that the extrapolation to zero field can be safely performed With other crystals containing more impurity, and consequently showing a smaller magneto-resistance effect, the measurements become unreliable at higher values of the field than with this one, and the extrapolation is accordingly longer Since, however, the variation with H^2 is correspondingly less rapid, the extrapolated value is about equally reliable

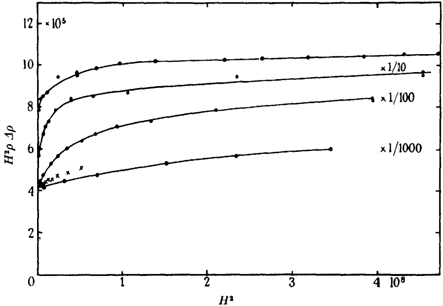


FIG 3—Magneto resistance effect in a long crystal of Bi + 0 0004 % Pb The horizontal scale of successive curves is enlarged by the factors shown The crosses denote points obtained using the Helmholtz coil to produce the small fields

5—DISCUSSION OF MAGNETO-RESISTANCE EFFECT

As an explanation of the observations recorded above, it is suggested that we must take account of the fact that we have a non-uniform distribution of the impurity atoms throughout the bismuth Many of the alloys used contained less than 0·1 atomic % of impurity, and often considerably less, even if this amount were distributed through the sample with perfect regularity, each impurity atom would be 10 atomic distances away from the next Since there are certain to be fluctuations in the density of foreign atoms—at least statistical, if no worse—it is obviously only a first approximation to treat such a medium as homogeneous Considering the problem from a classical viewpoint, the first effect of this inhomogeneity in the crystal will be to produce a non-uniform distribution of the current flowing through it If now the crystal is placed in a magnetic field, the extent of the departure from uniformity will grow rapidly greater as the field strength is increased This we know since the *averaged* affect of the impurities on the resistance over the whole specimen increases rapidly with the field Thus as the field is increased, not only does the resistivity of the specimen alter, but the current distribution in it changes also

The point will perhaps be made clearer by considering a very simple model Suppose we have two strips of bismuth one pure and the other impure, lying side by side, and we allow them to carry the total current "in parallel" We will suppose that each obeys the law expressed by equation (1) but with different values of the constants (the pure strip having $z = 0$), and we will proceed to calculate the resistance of the compound strip as a function of H Let A_1 and A_2 be the cross-sectional areas of the pure and impure strips respectively, and let the quantities B, C and z be suffixed 1 and 2 to correspond Then it is merely a matter of simple algebra to show that the magneto-resistance effect for the compound strip, written in a form to compare with equation (2) is given by

$$\frac{H^2\rho}{\Delta\rho} = \frac{1+\{B_2+C_2z_2^2+\alpha(B_1-B_2)\}H^2+\alpha B_1C_2z_2^2H^4}{\{B_1+\alpha(B_2-B_1)\}+\{B_1B_2+B_1C_2z_2^2(1-\alpha)\}H^2},$$

where α is written for the quantity $A_2/(A_1+A_2)$

The expressions for B and Cz^2 as given in the paper by Jones, show that the latter is always considerably smaller than the former—a fact which is borne out by the experimental values shown in Table I Accordingly we can neglect terms in Cz^2 with respect to those in B, and thus obtain

$$\frac{H^2\rho}{\Delta\rho} = \frac{1+\{B_2+\alpha(B_1-B_2)\}H^2+\alpha B_1C_2z_2^2H^4}{\{B_1+\alpha(B_2-B_1)\}+B_1B_2H^2} \tag{3}$$

Writing y for $H^2\rho/\Delta\rho$ and x for H^2 as before, this is seen to be of the form

$$y = \frac{1+Px+Qx^2}{R+Sx},$$

and such an equation can always be fitted fairly well to the experimental points Fig 4 shows some examples of the kind of agreement obtained by fitting such a curve at the four points $x(=H^2) = 0$, 5×10^6, 5×10^7, 5×10^8 oersteds² The agreement is good, but not perfect, as is shown by the fact that rather different values are obtained for the quantities P, Q, R, S if the

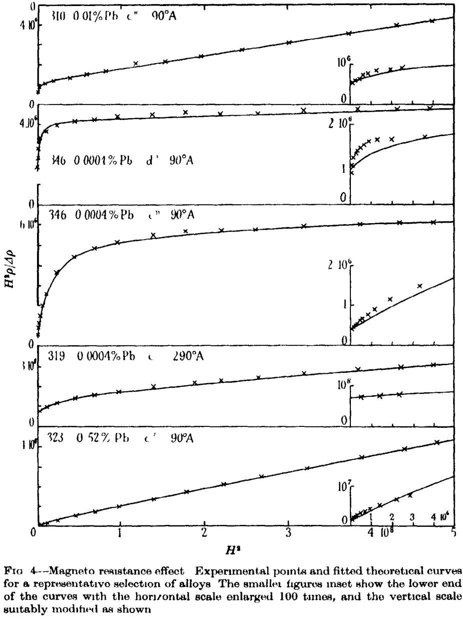


FIG 4—Magneto resistance effect Experimental points and fitted theoretical curves for a representative selection of alloys The smaller figures inset show the lower end of the curves with the horizontal scale enlarged 100 times, and the vertical scale suitably modified as shown

equation is made to fit at four other points Considering the crudity of the model used, however, the success is quite surprising *

Using the values of P, Q, R, S obtained in this way, we can solve for B_1, B_2, α and $C_2z_2^2$ This has been done for a series of crystals covering a range of different impurity contents The current was flowing parallel to the principal axis (case "*c*"), and the data all refer to a temperature of 90° abs The results are shown in Table I It is interesting to note that while B_1 remains approximately constant until a considerable amount of impurity is present, B_2 shows a steady decrease, and α and $C_2z_2^2$ a steady increase as the amount of impurity is increased This is just what would have been expected, but it does not seem justifiable to pursue the matter into more detail It is not, of course, suggested that the model used for the calculations gives any accurate picture of the real state of affairs in these very dilute solutions, or even that anything actually takes place which could be called a segregation of the impurity, as distinct from a mere random distribution of a very small proportion of foreign atoms For instance, one crystal was annealed for 3 days at a temperature only one or two degrees below its melting-point, and the results after annealing differed only slightly from those before

TABLE I

Sp no	% imp	$B_1 \times 10^6$	$B_2 \times 10^6$	α	$C_2z_2^2 \times 10^8$
316	0 0004 Pb	2 73	0 35	0 105	0 12
319	0 01	2 55	0 31	0 276	0 71
321	0 104	4 91	0 045	0 39	3 84
325	0 21	4 89	0 02	0 59	2 75
323	0 51	2 5	0 039	0 68	4 71
303	1 0	0 63	0 008	0 85	1 25
341	0 0003 Se†	3 04	0 42	0 347	0 0144
338	0 0008	2 59	0 37	0 391	0 0425
336	0 002	6 10	0 36	0 627	0 099
334	0 01	1 84	0 10	0 558	0 375
350	0 104	0 525	0 0047	0 870	0 366

† These, and all other selenium contents, are to be considered as added to the 0 0004 % of lead present in the bismuth as supplied by the makers

	B_1	B_2	α	$C_2z_2^2$
Before	$2\ 55 \times 10^{-6}$	$0\ 31 \times 10^{-6}$	0 276	$0\ 71 \times 10^{-8}$
After	$2\ 82 \times 10^{-6}$	$0\ 38 \times 10^{-6}$	0 34	$0\ 47 \times 10^{-8}$

---

* It can further be shown that if instead of two, we take n strips in parallel, each with a different impurity content, we obtain an expression of the form

$$y = \frac{1 + a_1x + a_2x^2 + \quad + a_nx^n}{b + b_1x + \quad + b_{n-1}x^{n-1}}$$

giving $2n$ arbitrary constants, and enabling a curve to be fitted at $2n$ points

The annealing was carried out *in vacuo*, and the change in $C_2z_2^2$, the only significant alteration, is explicable as due to the loss of some of the lead by volatilization. Table II shows the variation of the four quantities, referred to one specimen as the temperature is varied. The steady increase in α as the temperature is raised is rather puzzling, since it might have been expected that it would remain constant while the rest varied. It might be interpreted as meaning that the disturbance caused by the presence of an impurity atom extends to greater distances through the lattice at higher temperatures.

TABLE II.—SPECIMEN 319. 0.01% PB. *I* PARALLEL TO *oz*

T° abs	$B_1 \times 10^6$	$B_2 \times 10^6$	α	$C_2z_2^2 \times 10^8$
66	3.72	0.32	0.08	3.33
90	2.55	0.31	0.276	0.71
168	0.55	0.04	0.65	0.078
230	0.26	0.015	0.83	0.20
292	0.096	0.005	0.90	0.18

6. HALL EFFECT

Turning now to the Hall effect we first of all present in figs. 5*a* and *b*, two groups of curves summarizing the results obtained for the Hall coefficient R, defined by

$$E = \frac{RHI}{d},$$

where H = magnetic field, I = current, E = observed e.m.f., d = thickness of specimen, all in absolute e.m.u. Two different orientations of the crystal were considered.

1—I parallel to oz, H parallel to oy, E perpendicular to both I and H $(= R_\perp)$.

2—I parallel to oy, H parallel to oz, E perpendicular to both I and H $(= R_\parallel)$.

We see at once, in view of what has been written above, that it is very difficult to draw any unambiguous conclusions from these experimental data. The calculation of the Hall coefficient of a non-homogeneous medium, even on the simplest possible model, presents difficulties which render it not worth while in the present context. One point, however, which emerged before these difficulties were realized, seems worth recording Jones' theory leads to the following expression for the Hall coefficient

$$R = \frac{A - \left(\frac{c\Omega}{e}\right) zCH^2}{1 + z^2CH^2}, \tag{4}$$

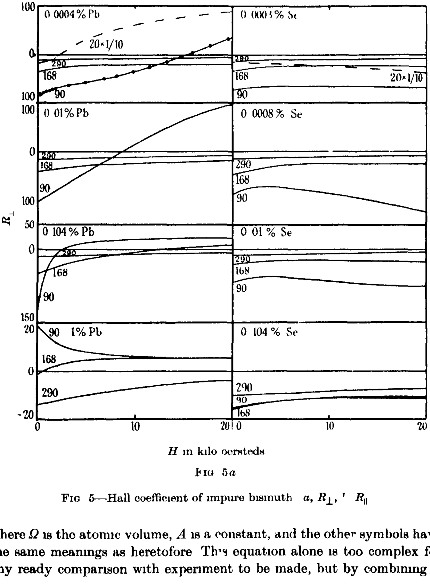


Fig 5*a*

Fig 5—Hall coefficient of impure bismuth *a*, $R_\perp$, ' $R_\parallel$

where Ω is the atomic volume, A is a constant, and the other symbols have the same meanings as heretofore This equation alone is too complex for any ready comparison with experiment to be made, but by combining it with the magneto-resistance equation (1) we easily obtain the following form

$$\frac{RH^2\rho}{\Delta\rho} = \frac{A - \left(\frac{c\Omega}{e}\right) zCH^2}{B} \tag{5}$$

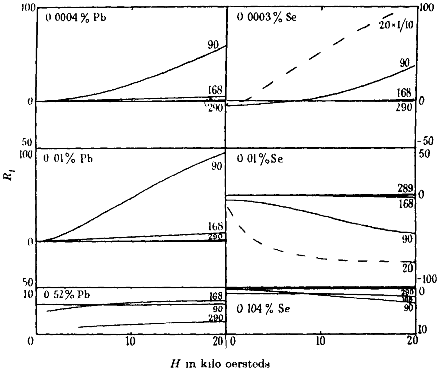


FIG 5b

Showing that the experimental quantity $RH^2\rho/\Delta\rho$ should be a linear function of H^2 It was in fact found that, with the same limitations as set forth on p 31 a very good straight line was obtained, as is shown by fig 6 It can hardly be fortuitous that just the particular and rather peculiar combination of observables predicted by the theory should give such an excellent fit, and that, moreover, under just those conditions that give the best agreement in the case of the magneto-resistance results

For brevity, let us refer to these conditions, expressible analytically by $z^2CH^2 \gg 1$, as "saturation" Returning for the moment to equation (3), we see that the slope of the curve of y vs x at saturation is $\alpha C_2 z_2^2/B_2$ whereas the original equation (2) gives a slope of Cz^2/B The change corresponds to the fact that at saturation all the current is flowing through that part of the compound strip to which we have given the suffix 2 If we attempt to calculate the Hall coefficient of such a compound strip at saturation, it is clear that the result depends not only on the relative size of the two parts, as does the magneto-resistance coefficient, but also on their shape If, for example, the two strips lie side by side in the plane normal to H, then at saturation,

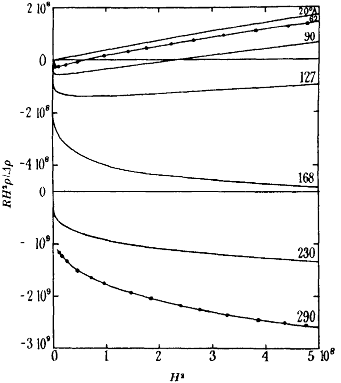


FIG 6—Hall and magneto resistance effects Experimental curves showing $RH^2\rho/\Delta\rho$ plotted against H^2 Bi + 0 0004 % Pb, 20–290° abs

when all the current is flowing through one of them, the Hall e m f across it is independent of its size If, on the other hand, the two strips lie one above the other when viewed in the direction of H, the Hall e m f across one at saturation (the other being treated as an insulator) is inversely proportional to its thickness in fact is proportional to $1/\alpha$ Thus in actuality, taking some intermediate case, we may expect to find that the Hall e m f 's at saturation are all too big, corresponding to the fact that the thickness of metal actually carrying the current is less than the thickness as measured with a screw gauge The ratio is not easily calculable, but might be expected to be of the order of the quantity α

Now we see that at saturation, (4) degenerates into

$$R_\infty = -\frac{c\Omega}{ez}, \tag{6}$$

the Hall coefficient being independent of both field and temperature This, being true for all values of the impurity content, will also be true for the compound strip, provided we introduce the constant factor $\approx\alpha$ mentioned above That this conclusion is supported by the experiments, is shown very clearly by figs 5a and b For the purest bismuth (z small) even at hydrogen temperatures (C large) it was impossible to produce sufficiently high fields to reach saturation although, as will be seen, we are approaching that condition For the more impure crystals, saturation could be obtained at temperatures as high as 168° abs The values of the constant Hall coefficients so obtained, give us, according to equation (6), an estimate of z Table III shows the values of z calculated so, compared with the known impurity contents The ratio of the two is the above-mentioned factor $\approx\alpha$, as will be seen, it is of the right order of magnitude, albeit again showing inexplicable systematic variations

TABLE III

R	z (calc)	z (actual)	Ratio
	$R_\perp$		
5 20	0 0425	1 0	0 042
8 10	0 0272	0 52	0 052
13 0	0 0170	0 21	0 081
21 0	0 0105	0 104	0 10
800	0 000275	0 0004 Pb	0 67
− 90	− 0 00245	− 0 01 Se	0 25
− 11	− 0 0200	− 0 104	0 19
	$R_\parallel$		
6 20	0 0356	0 52	0 069
625	0 00035	0 0004 Pb	0 88
− 78 0	− 0 00284	0 01 Se	0 28
− 5 (approx)	− 0 044	0 104	0 42

7—THE GALVANO-MAGNETIC CONSTANTS OF PURE BISMUTH

It was hoped when this investigation was commenced that, by proceeding to the limit as the amount of impurity was progressively decreased, an estimate could be made of the galvano-magnetic constants of really pure bismuth. The considerations put forward in the preceding sections have made this more difficult than was anticipated, and the results have not been as satisfactory as those already presented. Even with the purest sample available, the magneto-resistance coefficient still showed considerable variation with field in the region of small H, whereas the theory indicates

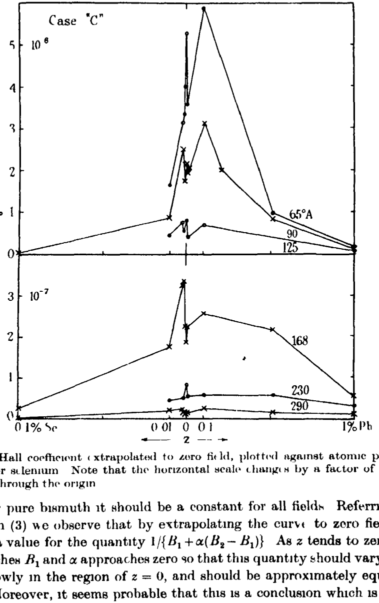


FIG 7—Hall coefficient extrapolated to zero field, plotted against atomic percent of lead or selenium. Note that the horizontal scale changes by a factor of ten in passing through the origin.

that for pure bismuth it should be a constant for all fields. Referring to equation (3) we observe that by extrapolating the curve to zero field we obtain a value for the quantity $1/\{B_1+\alpha(B_2-B_1)\}$. As z tends to zero, B_2 approaches B_1 and α approaches zero so that this quantity should vary only very slowly in the region of $z = 0$, and should be approximately equal to $1/B_1$. Moreover, it seems probable that this is a conclusion which is inde pendent of the details of the "inhomogeneity" theory presented above This method of estimating B_1 has the additional advantage that effects due to differing orientations of the binary axes tend to vanish at sufficiently low values of the field It is unfortunate that it involves an extrapolation where the experimental points are least accurate, as this seriously impairs the accuracy of the results However, fig 7 shows the values of B_1 for zero field ($= B_0$) obtained in this manner and corrected for the small variations in temperature of the original observations, plotted against the impurity content, for case "*c*" The variations, although large, are random, and not inconsistent with the errors of the extrapolation The results for cases "*d*" and "*e*" are essentially similar

It is interesting to compare with these fig 8 which shows the magneto-

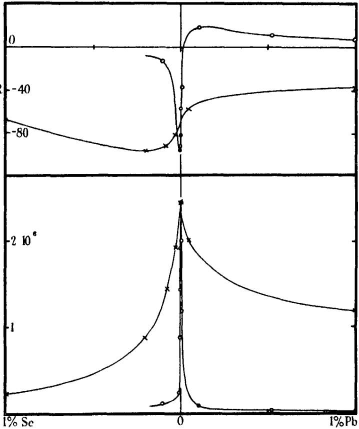


FIG 8 o——o Hall coefficient (above) and magneto resistance coefficient (below) at $T = 76°$ abs, $H = 5000$ oersteds, plotted against impurity content ×——× The same data, but with the horizontal scale enlarged 100 times resistance coefficient for the arbitrary values $T = 76°$ abs , $H = 5000$ oersteds, also plotted against the percentage impurity The rate of variation with z is seen to be quite a different order of magnitude The point shown for the value $z = 0$ is of particular interest, as it is taken from some recent results of de Haas, Gerritsen and Capel (1936) on bismuth which they had carefully purified by repeated recrystallizations The point is in excellent agreement with our own results, and it would be interesting to know how this specimen behaved in very small fields The figure shows very clearly the enormous changes which are produced at high fields and low temperatures by minute quantities of impurity This result is of some importance, in view of the common practice of using the magneto-resistance effect to measure magnetic fields It suggests that to obtain the greatest sensitivity it would be worth while using the purest bismuth

Fig 8 also brings out very well yet another point The upper part of the diagram shows the values of the Hall coefficient for the same temperature and field strength, as that to which the magneto resistance data refer Comparing the two curves with the extended scale of z it is very clear how in the neighbourhood of $z = 0$ the magneto-resistance effect depends on z^2 while the Hall effect is determined by z—in agreement with the prediction of Jones' theory

From fig 7 and corresponding diagrams for "d" and "e" we estimate the following as probable values for B for pure bismuth for the three different orientations

TABLE IV

$T°$ abs	$B_c \times 10^8$	$B_d \times 10^8$	$B_e \times 10^8$
20	6000	2500	1500
65	420	120	550
90	210	50	250
125	65	15	70
168	25	4 5	23
230	6	1 1	5 5
290	2	0 28	1 4

We next refer to equations (19)–(21) of Jones' paper, giving theoretical expressions for the same quantities These can be written in the form

$$\left.\begin{aligned} (B_c/\sigma_{\parallel}\sigma_{\perp}) &= x_{\perp}(1-x_{\perp})\ (c/en_B)^2, \\ (B_d/\sigma_{\perp}^2) &= x_{\perp}(1-x_{\perp})\ (c/en_B)^2, \\ (B_e/\sigma_{\parallel}\sigma_{\perp}) &= x_{\parallel}(1-x_{\parallel})\ (c/en_B)^2, \end{aligned}\right\} \qquad (7)$$

where n_B is the number of free electrons/c c , $\sigma_{\parallel}$ and $\sigma_{\perp}$ the conductivities respectively parallel and perpendicular to the axis, and $x_{\parallel}$, $x_{\perp}$ the fraction of these conductivities due to the electrons alone, the remainder being due to "positive holes" The left-hand side comprises only quantities that are directly measurable, and thus graphs can be drawn showing the variation of the right-hand side with temperature (fig 9) The three curves are seen to

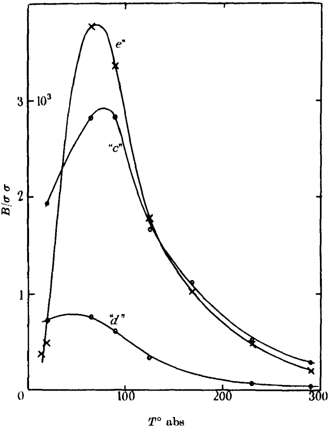


FIG 9—(See text) Note that curve "e" involving $x_{\|}(1-x_{\|})$ shows a much sharper fall below its maximum than "c" and "d" which both depend on $x_{\perp}(1-x_{\perp})$

be all of the same shape, and since n_B is a constant it is natural to associate the maximum with the maximum of the function $x(1-x)$, as x passes through the value 0·5 With this assumption we can calculate from the heights of the maxima the value of n_B, and hence the number of free electrons per atom The results for n_B are

Case "c", $n_B = 5·8 \times 10^{17}$ electrons/c c
Case "d" $11·2 \times 10^{17}$,,
Case "e" $5·0 \times 10^{17}$,,

giving a mean of $7\ 3 \times 10^{17}$, and taking the atomic volume as $3\ 53 \times 10^{-23}$ c c / atom, this gives the result of $2\ 6 \times 10^{-5}$ free electrons per atom No explanation is yet forthcoming of why the three different orientations lead to three consistently different values of n_B, but it should be noted that they differ only by a factor of about 2, being all of the order of 10^{-5} electrons per atom While it has been suspected for some time that this quantity would be very small for bismuth, no direct determination has hitherto been available The result is consistent with previous estimates, however, and seems to establish beyond all doubt that the exceptional galvano magnetic properties of bismuth are due to its having so small a number of effective conduction electrons From the shape of the curves of fig 9 we could now easily calculate how $x_{\|}$ and $x_{\perp}$ vary with the temperature, except that there would be an ambiguity as to whether we have really calculated x or $(1-x)$ It was hoped that the Hall effect data would decide this point, and, proceeding by analogy with the magneto resistance effect, we take the values of R extrapolated to zero field as giving the best approximation to the values for pure bismuth The resulting figures for $R_{\perp}$ are shown plotted against impurity content in fig 10, and here it is not difficult to estimate what the value for pure bismuth would be For $R_{\|}$ the results are rather surprising, in that for the purer specimens the Hall coefficient in small fields was too small to be measured (compare figs 5*a* and *b*) The rather sketchy results available are shown in Table V

TABLF V

%	65° abs	90 abs	125° abs	168° abs	230° abs	290° abs
0 52 Pb	8	6 5	6 5 (?)	4 8	3 4	1 8
0 21	—	7 5	—	4	—	0 7
0 104	—	4 (?)	—	—	—	1
0 01	<0	0	2 (?)	4	≈ 2	≈ 1
0 0004	<0	<0 (?)	>0	>0	>0	>0
0 0003	-13	-7	-5	$-2\ 5$	$-1\ 5$	<0
0 0008	—	-3	—	—	—	0 5
0 002	—	-2	—	—	—	1
0 01	$-8\ 5$	-7	-3	$-0\ 5$	1	0 7
0 104 Se	—	$-0\ 5$	—	$-1\ 5$	—	<0 (?)

This result is believed to be new, as nobody has previously made any measurements in low fields The interpolated values of both $R_{\|}$ and $R_{\perp}$ for pure bismuth are given in Table VI

The relevant equations from Jones' paper giving theoretical expressions for $R_{\|}$ and $R_{\perp}$ in terms of n_B, $x_{\perp}$ and $x_{\|}$ are

$$\left.\begin{aligned} R_{\|} &= (2x_{\perp}-1)\ \ (c/en_B), \\ R_{\perp} &= (x_{\perp}+x_{\|}-1)\ \ (c/en_B), \end{aligned}\right\} \qquad (8)$$

and since we must have $0 \leqslant x \leqslant 1$, we find that the greatest possible value of en_B/c is 0·027, corresponding to $6{\cdot}0 \times 10^{-5}$ electrons per atom This is of the same order of magnitude as the value deduced from the magneto-resistance data, but it leads to the conclusion that $x_\perp = 0{\cdot}5$ for all temperatures,

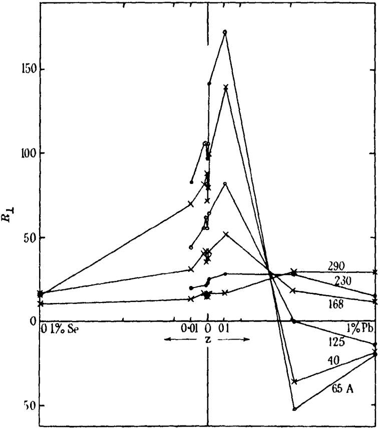


FIG 10—Hall coefficient ($R_\perp$) extrapolated to zero field plotted against impurity content The horizontal scale again changes by a factor of ten at the origin

TABLE VI

T° abs	$R_\perp$	$R_\parallel$
20	180 (approx)	—
65	103	−6
90	80	−3
125	57	−2
168	38	−1
230	22	≈0
290	15	0

while x falls continuously from 1 0 at 20° abs to 0·55 at 290° abs Thus it is seen that while there is general agreement as to the order of magnitude of n_B, the two groups of data are not consistent as to the temperature variation of the x's Nevertheless there has been a sufficiently striking agreement between theory and the experimental results presented in the earlier part of this account to indicate that the final solution must be sought on the lines here laid down

It is a pleasure to acknowledge my indebtedness to the Department of Scientific and Industrial Research for a personal grant during the tenure of which this work was done, to Dr H Jones for many helpful discussions on the theoretical aspects of the subject, and to Professor A M Tyndall for his continued interest in the progress of the work

REFERENCES

Becquerel, J 1912 *C R Acad Sci Paris* **154**, 1795–98
Borelius G and Lindh, A E 1917 *Ann Phys*, *Lpz*, iv, **53**, 97–137
Campbell, L L 1923 'Galvanomagnetic and Thermomagnetic Effects ' London Longmans Green and Co
van Everdingen, E 1897 *Commun Phys Lab Leiden*, **37**, 61
Gruneisen, E and Gielessen, J 1936 *Ann Phys*, *Lpz*, v **26**, 449–64
— — 1937 *Ann Phys Lpz*, **28**, 225–39
de Haas W J, Gerritsen, A N and Capel, W H 1936 *Physica*, *'s Grav*, **3**, 1143 58
Heaps, C W 1927 *Phys Rev* **30**, 61–5
Jones, H 1936 *Proc Roy Soc* A **155**, 653–63
Kapitza, P 1928 *Proc Roy Soc* A, **119**, 401–43
Lownds, L 1901 *Ann Phys*, *Lpz* iv, **6**, 146–62
— 1902 *Ann Phys*, *Lpz*, iv **9** 677–90
Onnes, H K and Beckman, B 1912 *Commun Phys Lab Leiden*, **129**, 132
Schubnikov, L and de Haas, W J 1930 *Commun Phys Lab Leiden* **207**
Stierstadt, O 1933*a* *Z Phys* **80** 636–65
— 1933*b* *Z Phys* **85**, 310–31
— 1933*c* *Z Phys* **85**, 697 707
– 1934 *Z Phys* **87**, 687–99
Thompson, N 1936 *Proc Roy Soc* A **155**, 111–23

Hyperfine Structure and Nuclear Moments of Aluminium

By D A Jackson and H Kuhn

Clarendon Laboratory, Oxford

(*Communicated by F A Lindemann, F R S—Received* 28 *July* 1937)

[Plate 1]

Introduction

The structure of the resonance lines of the arc spectrum of aluminium was investigated by the method of absorption in an atomic beam, and the structure of certain other lines of aluminium, emitted by a liquid-air-cooled hollow cathode tube, was also investigated The nuclear spin of aluminium was thus shown to be $\frac{5}{2}$ and the magnetic moment about 4 0 nuclear magnetons A note (Jackson and Kuhn 1937*a*) containing these results was published elsewhere

The light source and the spectroscope

The arc spectrum of aluminium possesses two groups of resonance lines,

$$3\,{}^2P_{\frac{1}{2},\frac{3}{2}}\text{–}4S_{\frac{1}{2}}\ (3962\text{ A},\ 3944\text{ A})$$

and $$3\,{}^2P_{\frac{1}{2},\frac{3}{2}}\text{–}3\,{}^2D_{\frac{3}{2},\frac{5}{2}}\ (3092\ 7\text{ A},\ 3092\ 8\text{ A},\ 3082\text{ A})$$

The source of these lines was a fused silica discharge tube, excited by external electrodes connected to a ½ kW high-frequency oscillator, the ends of the discharge tube were 8 cm long and 5 cm in diameter, and the capillary was 5 cm long and 6 mm in internal diameter The tube was filled with neon at 3 mm pressure, and the capillary was fitted with a side tube of 5 mm internal diameter and 10 cm length containing aluminium tribromide When this side tube was heated by being immersed in a Dewar vessel containing water at 45° C, the discharge tube on excitation emitted the aluminium resonance lines very strongly, the neon lines being greatly weakened With this light source the aluminium lines could be photographed in 10 sec and they were entirely free from self-reversal In order to obtain these lines equally free from self-reversal in a liquid-air-cooled hollow cathode tube an exposure thirty times longer was needed

For the atomic beam experiment, the spectroscope ($f = 150$ cm) described in earlier work was used, also the etalon was the same

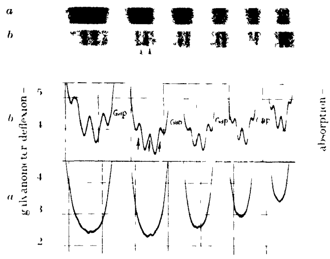


Fig. 2. $3\,P_{\frac{1}{2}} - 4\,S_{\frac{1}{2}}$

(*a*) light source (*b*) absorption by atomic beam

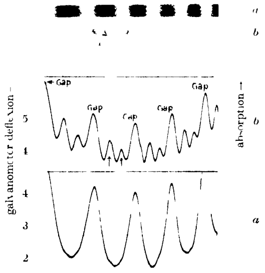


Fig. 3. $3\,P_{\frac{1}{2}} - 3\,D_{\frac{3}{2}}$

(*a*) light source (*b*) absorption by atomic beam

The production of the atomic beam

It was impossible to produce an atomic beam of aluminium by evaporating aluminium in a silica tube in the manner used for silver, for two reasons first, silica is very rapidly attacked by aluminium, and secondly, it softens at a temperature below that at which aluminium has a sufficiently high vapour pressure However, on account of the very great intensity of the resonance lamp, an atomic beam was only needed for a short time, the length of exposure being 10–15 sec This was first achieved by evaporating aluminium from a helix of tungsten wire of 1 mm diameter, the helix consisting of 7 turns of 1 cm diameter and 7 mm pitch, on each turn was closely wrapped 10 mm of 1 mm diameter aluminium wire It was possible to evaporate this in about 30 sec , and by placing a slit at a suitable distance above the helix an atomic beam was made It was also possible to observe the structure of the resonance lines by this arrangement, but the supply of aluminium for the atomic beam was of very short duration and rather inadequate

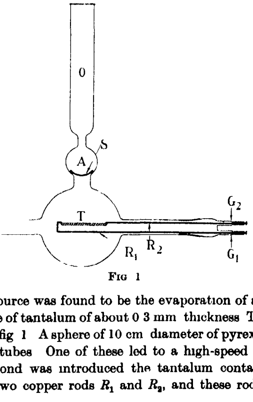


Fig 1

A far better source was found to be the evaporation of aluminium from a container made of tantalum of about 0 3 mm thickness The arrangement used is shown in fig 1 A sphere of 10 cm diameter of pyrex glass was fitted with three side tubes One of these led to a high-speed mercury pump, through the second was introduced the tantalum container which was carried by the two copper rods R_1 and R_2, and these rods were screwed internally into the two copper-glass seals G_1 and G_2, the joint between the side tube and the glass tube with the copper-glass seals was made tight with black wax The third side tube led to the absorption chamber *A* This was a horizontal glass tube (with its axis in the line of sight of the spectrograph) of 3 cm diameter and 8 cm length, and at each end a fused silica window was fixed on with sealing wax The evaporating aluminium reached this chamber through the slit *S* in a nickel plate which entirely covered the opening into the absorption chamber This slit varied in width from 14 to 3 mm , according to the degree of collimation required At the top of the absorption chamber was fitted a tube *O* of 3 cm diameter and 20 cm length with a window at the upper end, and through this the tantalum container *T* could be observed, the aluminium deposit here being far less than on the walls of the sphere

The tantalum container was about 5 cm in length Its cross-section was V-shaped, each side being 5 mm and the opening at the top about 4 mm The distance from the tantalum container to the slit was about 7 cm The tantalum container and the slit were parallel to each other, and perpendicular to the line of sight of the spectrograph The greatest slit width corresponded to a collimation of 7 1 and the smallest about 20 1 The entire sphere was immersed in water for cooling The tantalum container held four pieces of aluminium wire 15 mm long and 2 mm in diameter With a current of about 80 amp these were evaporated in about 90 sec the rate of evaporation being fairly uniform

The current required to produce the necessary rate of evaporation was determined experimentally and was accurately reproducible, as the size of the tantalum container and the quantity of aluminium used were always repeated exactly It was found that by keeping the watts constant a nearly constant rate of evaporation could be maintained This was ascertained by taking three or four consecutive photographs of 10 sec exposure Two or three of these were found always to have nearly equal absorption A still better constancy of rate of evaporation was attained by gradually reducing the watts of the heating current throughout evaporation, and in this way it was possible to make four to six photographs of the absorption with one charge of aluminium The appropriate variation of the current was determined experimentally The necessity of reducing the current arose from the increased resistance of the tantalum container as the quantity of conducting aluminium decreased

Though the tantalum containers were not much attacked by the aluminium, a new container was used for each evaporation If a container was used twice, the conditions of evaporation were not reproducible

The degree of vacuum obtained in the sphere during evaporation of the aluminium was extremely high The window at the top of the observation tube was about 23 cm from the slit *S* The "shadows" of the edge of the slit could, nevertheless, very clearly be seen on the window The perfection of the vacuum was undoubtedly due to the "gettering" effect of the evaporating aluminium

THE STRUCTURE OF THE LINE 3 $^2P_{\frac{1}{2}}$–4$S_{\frac{1}{2}}$ λ 3944 A

This line was found to possess three components, the structure being observed with etalons of 2 and 3 cm plate separation and 1 12 collimation of the atomic beam * The plates made with an etalon of 3 cm separation were measured visually on a micrometer, and the mean result for the positions of the three components was −0 047, 0 000 and +0 047 cm $^{-1}$ Those made with an etalon of 2 cm separation were measured both visually and also by measuring the positions of the minima on the photometer traces The results obtained by these methods were −0 048, 0 000 and 0 048 cm $^{-1}$, the probable error being between ±0 001 and ±0 002 cm $^{-1}$ The positions of the three components are thus −0 048 0 000 and +0 048 cm $^{-1}$, and the probable error ±0 001 cm $^{-1}$

The intensity of the three components appeared very nearly equal A photograph of the line with and without absorption, together with photometer traces, is shown in fig 2, Plate 1 The absorption lines, in one order are marked by arrows

The intensity of the components can be determined by photographing first the light source only, and then the light source with absorption due to the atomic beam An intensity scale is put on the plate, and a photometer trace of the plate is made The ratio of the absorption coefficients (the difference in the logarithm of the intensity at the position of the absorption in the photograph with absorption and that without absorption) for the various components gives at once the true intensity ratio This has already been done with a very high degree of accuracy for silver (Jackson and Kuhn 1937*b*) The principal condition for accuracy is constancy of the light source, and this condition was very perfectly fulfilled by the resonance lamp used Test plates made showed that provided the lamp was run for 2 min to reach thermal equilibrium, photographs of 15 sec exposure taken at intervals of 5 min were found to be of equal intensity to within 2 %, any error being due not to variation in light but to uncertainties in length of exposure, which could amount to $\frac{1}{5}$ sec

---

* The collimation of the atomic beam is the ratio of half the sum of width of slit and width of aluminium container to the distance between them, this is approximately equal to the ratio of the velocity in the line of sight to the total velocity of the atoms

For the determination of the nuclear spin, the intensity ratio of the central component to the outer components (which are of equal intensity) has to be measured It was, however, found impossible to get accurately consistent values The reason can easily be found in the entirely different conditions of the background for the outer and inner components The outer absorption lines fall on the edge of the emission line, where the intensity curve is falling very steeply This leads to only a small error in the measuring of the distance of these lines (about 2 %), but to a very considerable error in the determination of the intensity of the background and therefore of the absorption coefficients, it makes the absorption due to the outer components appear too weak Moreover, a purely erratic error in determination of the position in the background corresponding to the absorption by the outer components causes the background to appear systematically weaker, for the reduction due to an error outwards is much greater than the increase due to an equal error inwards

Six plates were measured, three fringes being measured in each The results varied between 1 04 1 and 1 23 1 for the ratio of the intensity of inner component to that of the outer components But as these measurements are liable to the above-described systematic error, tending to increase the ratio, it is clear that the value must be substantially less than 1 2 1 This indicates a high value for the nuclear spin, the intensity ratio for a spin of $\frac{3}{2}$ being 1 2 1 and that for an infinitely great spin 1 1 It would have been possible to eliminate the systematic error by broadening the emission line But even if this were done, the intensities measured would not have been of much value, on account of the very small change in intensity ratio for a change in spin value

THE STRUCTURE OF THE LINE $3\,^2P_{\frac{1}{2}}$–$4S_{\frac{1}{2}}$, 3962 A

The structure of this line was examined with etalons of plate separations 2, 3 and 4 cm The atomic beam gave an absorption line with a width of 0 04 cm $^{-1}$, which could not be resolved into components even with the greatest collimation (20 1) and etalon length Under these conditions it would have been possible to resolve lines with separations of about 0 012 cm $^{-1}$, it appears, therefore, that this line must possess more than four components (probably six, which is the greatest number which a line of this type can possess)

THE STRUCTURE OF THE LINE $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$, λ 3082 A

This line was investigated by means of an atomic beam with a 2 cm etalon and was found to possess two components (fig 3, Plate 1) The absorption lines are marked by arrows The separation of these was measured both directly and also by the measurement of photometer traces, the first method giving the value 0 066 ± 0 002 cm $^{-1}$ and the second 0 067 ± 0 002 cm $^{-1}$ The structure of the lines was observed with collimations of the atomic beam of 7 1 and 12 1 The intensity of the component of longer wave-length was the greater It was possible to measure the intensity ratio (ratio of absorption coefficients) of the two components to a very high degree of accuracy The absorption lines lay well within the emission line and were placed symmetrically on either side of the maximum, so that there was no systematic error of the type which made impossible the accurate determination of the intensities of the three components of the line $3\,^2P_{\frac{1}{2}}$–$3S_{\frac{1}{2}}$ The intensity of the atomic beam required to produce absorption was considerably less than that needed for the line $3\,^2P_{\frac{1}{2}}$–$3S_{\frac{1}{2}}$ and the exposure needed was only 10 sec, consequently it was possible without difficulty to make four successive photographs of the absorption with one charge of aluminium in the tantalum container A considerable source of error was the purely erratic error inherent to all photographic intensity measurements In the ultra-violet, where the contrast of plates is relatively low, this introduces a possible error of about 4 % It is, however, doubled in the method of determining intensities by absorption, for it is here necessary to determine the ratio of the logarithms of two intensity ratios Any one determination of the intensity ratio of the two components could therefore be expected to be liable to an erratic error up to about 8 % In order to reduce this erratic error a large number of intensity measurements was made

The method was exactly similar to that used by the authors in their determination of the intensity ratios in the hyperfine structure of the resonance lines of silver A cross wire was fitted to one end of the slit A photograph of the light source (exposure 10 sec) was made, then four exposures (10 sec) of the light source with absorption by the atomic beam, and finally a series of intensity marks was made by giving a succession of exposures (10 sec) to a hydrogen tube with various slit widths Photometer traces were made, and by reference to the cross-wire the exact position was found of the absorption lines, in the emission line without absorption Only those plates were measured in which the degree of absorption in consecutive exposures was constant to within about 30 % (this showing that the density of the atomic beam was fairly constant). In each plate between two and five fringes were measurable, giving two to five intensity determinations. The absorption coefficient of the stronger line ranged between 0.18 and 0.47 (expressed as $\log_{10}$) in the various photographs measured. Two series of measurements were made, the first with a collimation of the atomic beam of 7 : 1 and the second 12 : 1. The results of these are given in

TABLE I—COLLIMATION OF ATOMIC BEAM 7 : 1

Absorption coefficient		Individual determinations of intensity ratio					Mean
A_3	0.22	1.28	1.13	1.15			1.19
A_2	0.30	1.23	1.18	1.18	1.28		1.22
B_1	0.33	1.24	1.15	1.19			1.19
C_1	0.34	1.20	1.21	1.14	1.17		1.18
C_2	0.34	1.16	1.17	1.12	1.13		1.15
A_1	0.38	1.24	1.17	1.16	1.15		1.18
B_2	0.39	1.20	1.18				1.19
C_3	0.41	1.23	1.18	1.18	1.28		1.22
C_4	0.47	1.19	1.16	1.14	1.12	1.19	1.16

Mean of 33 measurements 1.18_4
Corrected for Doppler wing 1.20_6
Mean error 0.033, probable error of mean value 0.006

TABLE II—COLLIMATION OF ATOMIC BEAM 12 : 1

Absorption coefficient		Individual determinations of intensity ratio					Mean
A_1	0.18	1.34	1.21	1.18	1.24		1.24
B_2	0.28	1.20	1.21	1.16	1.14		1.18
A_2	0.30	1.24	1.23	1.21	1.24		1.23
B_1	0.36	1.20	1.21	1.17	1.17		1.19
C_1	0.35	1.14	1.17	1.20			1.17
C_2	0.34	1.19	1.13	1.16	1.22		1.18
A_4	0.42	1.22	1.25	1.17	1.15	1.23	1.20
A_3	0.44	1.20	1.24	1.23	1.19	1.17	1.21

Mean of 33 measurements 1.20_0
Mean error 0.035, probable error of mean value 0.006

Tables I and II. In these tables each row gives the measurements for any one exposure. In the first column is given the mean value of the absorption coefficient for the stronger line and the series to which each exposure belonged (thus A_1, A_2, A_3, A_4 are four consecutive exposures during which the atomic beam was running continuously), in the second the values of the intensity ratio of the doublet according to each fringe measured, and in the third the mean of these measurements. To the measurements made with an atomic beam collimation of 7 1 a small correction has to be applied on account of the overlapping of the maxima of each component by the Doppler wing of the other The total half-value width of the absorption lines is 0 027 cm $^{-1}$ The intensity of absorption due to each line is therefore equal, at the position of the other line, to $(\frac{1}{2})^{\frac{0\ 066}{0\ 014}} = 0\ 05$ of its maximum value From this it follows that the observed intensity ratio ι is given by

$$\iota = \frac{I + 0\ 05}{1 + 0\ 05\, I},$$

where I is the true intensity ratio, and according to this equation an observed intensity ratio $1\ 18_4$ is given by a true intensity ratio $1\ 20_6$ With an atomic beam collimation of 12 1 the corresponding correction is negligible

The possibility of systematic errors must be considered These can arise in three ways from unresolved structure due to the term $3\,^2D_{\frac{3}{2}}$, from insufficient resolving power of the etalon, and from alteration in the intensity of the atomic beam during the course of an exposure

The separation of the two observed components of the line $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$ was 0 018 cm $^{-1}$ greater than that of the two levels of the term $3\,^2P_{\frac{1}{2}}$ From this it follows (see p 59) that the total splitting of the term $3\,^2D_{\frac{3}{2}}$ is about 0 03 cm $^{-1}$, and that each of the observed lines has an unresolved structure of width about 0 02 cm $^{-1}$ that of the stronger line being somewhat greater This structure is less than the half-value width with a collimation of 7 1 (0 027 cm $^{-1}$) and more than that with a collimation of 12 1 (0 016 cm $^{-1}$) The possible influence of this structure would be to make the effective intensity of the stronger line somewhat smaller, on account of its being spread over a slightly greater width But this effect would be very much stronger with the higher degree of collimation The perfect agreement of the results obtained with the two collimations shows, therefore, that any error must be very small, probably not more than 0 01

The resolving power of the etalon was at this wave-length between 0 025 and 0 030 cm $^{-1}$ The width of the lines, even with the higher collimation, is rather greater than this, on account of the unresolved structure described above This width excludes the possibility of loss of apparent intensity due to too much concentration of the absorption compared with the resolution of the etalon

A further error arising from finite resolving power of the etalon is the intensity, at the positions of the absorption lines, due to the remaining emission intensity between and on either side of the absorption lines The absorption maxima have distances from the neighbouring emission peaks of about three times the instrumental half-width Here, the intensity has dropped to less than one-tenth of its maximum value This effect makes the intensity ratio appear smaller An approximate calculation shows that this error must be well under 0 01 A similar calculation applied to the authors' results for silver would indicate a correction of 0 3, the observed value 2 7 1 becoming 3 1, in agreement with the theoretical value

The last source of systematic error to be considered is the influence of alteration of the intensity of the atomic beam during the course of an exposure The amount of this can be calculated a change in absorption coefficient from 0 30 to 0 20 during an exposure results in a loss of rather less than 0 02 in the measured intensity ratio for an intensity ratio of 1 2 1, while a change from 0 45 to 0 35 similarly gives rise to a loss of rather more than 0 02 The change in the average value of the absorption coefficient from one exposure to the next of the same series gives a good estimate of the amount of change in intensity of the atomic beam during these two exposures In this way it can be seen that the average value of the change is 0 04, and from the figures above it follows that a correction of about 0 01 must be applied to the average value of the intensity ratio

The intensity ratio of the two hyperfine structure components of the line 3 $^2P_{\frac{1}{2}}$–3 $^2D_{\frac{3}{2}}$ is thus shown to be 1 21 1, all possible sources of systematic error were considered and allowed for, the erratic error is probably about ±0 01 and certainly less than ±0 02, on account of the great number of measurements made The value 1 21 1 can therefore be relied on to be accurate to ±0 02

THE STRUCTURE OF THE LINE 3 $^2P_{\frac{3}{2}}$–3$D_{\frac{5}{2}}$, λ 3092 7 A

The above line could not be resolved into components Its width was about 0 04 cm $^{-1}$ and the resolving power of the etalon (2 cm plate separation) with which it was examined was 0 03 cm $^{-1}$, the half value width of the line in the atomic beam (collimation 12 1) being 0 016 cm $^{-1}$ From this it can be assumed that this line possesses more than two components, for under these conditions a doublet structure would have been resolved

On account of its small intensity and closeness to the line 3092 7 the line 3 $^2P_{\frac{3}{2}}$–3 $^2D_{\frac{3}{2}}$ (3092 8) was not examined by means of absorption in an atomic beam

THE STRUCTURE OF THE LINES 4$S_{\frac{1}{2}}$–5 $^2P_{\frac{3}{2}}$, λ 6696 A
AND 4$S_{\frac{1}{2}}$–5 $^2P_{\frac{1}{2}}$, λ 6699 A

These lines were emitted by the resonance lamp, their intensity being about equal to that of the neighbouring neon lines Their structure was investigated with an etalon with plates silvered so that each plate transmitted 1½ %, the etalon being used in conjunction with the 1½ m spectrograph and a very dense glass prism

The separation of these two lines by the dispersion of the prism was only about ¼ mm It was therefore necessary to use a plate separation of the etalon of such length that the order of interference of the two lines would differ by $n + \frac{1}{2}$ orders This condition was satisfied by using a plate separation of 1 5 cm In order to make the Doppler width of the lines as small as possible the capillary was cooled with water at 50° C

Both of the lines were found to consist of doublets only just resolved, the separation being 0 05 cm $^{-1}$ The half-value width of the lines is 0 04 cm $^{-1}$ This doublet structure was observed by Ritschl but he gave no intensity ratio

The resolution of the lines was just enough for the photometer to show a definite minimum between them Accordingly four plates of the line were made, each being given intensity scales by means of a V shaped slit The exposure of the lines was 1 hr , and three exposures of 1 hr of the V-shaped slit with different intensities of the light were made Photometer traces were made of the plates, and it was found that in two or three fringes in each plate the maxima and minima were just sufficiently well defined in the stronger line, $4S_{\frac{1}{2}}$–$5\,{}^2P_{\frac{1}{2}}$, for a comparison of the intensities of the two components In all ten measurements were made, the values found being 1 17, 1 12, 1 19, 1 20, 1 17, 1 18 1 13, 1 19, 1 18, 1 14 The component of longer wave-length was the stronger, the intensity ratio from the figures above being 1 16 ± 0 04 This value requires the usual correction for the overlapping of the Doppler wings The intensity of these wings at the maxima of the other components is equal to $(\frac{1}{2})^{\overline{0\ 05}\,|\,0\ 02} = 0\ 18$, so the true intensity ratio I is given by

$$1\ 16 = \frac{I + 0\ 18}{1 + 0\ 18 \times I}$$

This gives $I = 1\ 23 \pm 0\ 06$ No great accuracy is claimed for this value, as the maxima and minima were only just visible in the photometer traces Also the correction applied for overlapping is very great

THE STRUCTURE OF THE LINE Al II, $3S_0$-$3\,{}^3P_1$, λ 2669 A

A hollow cathode tube, in which the cathode was made entirely of aluminium, was used as a light source for the above line The tube was cooled with liquid air and the current used was 200 mA The intensity of the line was investigated with the tube filled with neon, helium and argon It was found to be the strongest with neon, and the most favourable pressure was about 1 mm

The structure of the line was investigated with a reflexion echelon mounted in the manner previously described by Jackson (1930) The possible resolving power of the grating was about 1,300,000 (0 03 cm $^{-1}$) at this wave-length, for it possessed 25 plates of thickness 7 mm However, as the half value width of the lines of aluminium at liquid-air temperature is about 0 06 cm $^{-1}$ the slit width of the spectrograph was made equal to 0 06 mm , corresponding to a resolving power of about 600,000 (0 06 cm $^{-1}$) In order to reduce still more the necessary length of exposure the plate was set normal, instead of at the appropriate tilt of 55° This increases the intensity two or three times but only the line under investigation is sharply focused Under these conditions a rather underexposed photograph of the line was obtained with a 2 hr exposure, and a good one with 4½ hr exposure

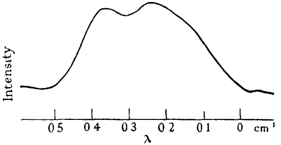


FIG 4—Al II, 2669 A

The photographs showed that the structure consists of a fine line of width about 0 1 cm $^{-1}$ with a much broader line, width about 0 2 cm $^{-1}$, of longer wave-length, the separation of the estimated centres of the two lines being 0 17 cm $^{-1}$ A photometer trace of this line is shown in fig 4

This observed structure can be satisfactorily explained as a triplet not completely resolved, the broad line being two lines separated by about 0 1 cm $^{-1}$, but not quite resolved, and the narrow line a single line The intensity distribution in the echelon spectrum makes the central component appear stronger, being nearest to the centre of the spectrum

THE STRUCTURES OF THE LINES $3s\ 3p^2\ {}^4P_{\frac{3}{2}}$–$3s\ 3p\ 4s\ {}^4P_{\frac{3}{2}}$, λ 3057 AND $3s\ 3p^2\ {}^4P_{\frac{5}{2}}$–$3s\ 3p\ 4s\ {}^4P_{\frac{5}{2}}$, λ 3050

The above lines were emitted by the liquid-air-cooled hollow cathode tube, and their structure was examined with the echelon grating referred to above (resolving power about 0 66 cm $^{-1}$) and also a small Lummer plate (resolving power about 0 03 cm $^{-1}$) The line 3057 appeared to possess two components separated by about 0 12 cm $^{-1}$, the width of each being about 0 1 cm $^{-1}$, a reduction in the current of the hollow cathode tube did not cause any reduction in the width of the components Under these conditions the Doppler width of the lines is about 0 05 cm $^{-1}$, it is therefore clear that the observed doublet structure is not a complete resolution, there must be finer unresolvable structure The line 3050 also appeared to possess two components, these were even wider than those of 3057, their width being about 0 14 cm $^{-1}$ and the separation of their centres about 0 16 cm $^{-1}$, this observed doublet structure is therefore also an incomplete resolution

THEORETICAL DISCUSSION OF RESULTS

From the observed triplet structure (-0 048 0 000 and $+0$ 048 cm $^{-1}$) of the line $3\,^2P_{\frac{1}{2}}$–$4S_{\frac{1}{2}}$ together with the doublet structure (0 00 and 0 05 cm $^{-1}$) of the line $4S_{\frac{1}{2}}$–$5\,^2P_{\frac{1}{2}}$, it can be seen that the levels $3\,^2P_{\frac{1}{2}}$ and $4S_{\frac{1}{2}}$ are both double with a splitting of 0 048 cm $^{-1}$ The structure of the line $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$ shows that the term $3\,^2D_{\frac{3}{2}}$ possesses an inverted structure the total width of which is about 0 03 cm $^{-1}$ This structure is probably due to perturbation by the term $3s\,3p^2\,^2D_{\frac{3}{2}}$ The splitting of the terms $4S_{\frac{1}{2}}$ and $3\,^2P_{\frac{1}{2}}$ can be explained as a normal hyperfine structure splitting The value of the nuclear spin can be calculated from the intensity ratios of the three components of the line $3\,^2P_{\frac{1}{2}}$–$4S_{\frac{1}{2}}$ and from the intensity ratio of the two components of the lines $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$ and $3S_{\frac{1}{2}}$–$5\,^2P_{\frac{3}{2}}$ The first of these was not measured to a very high degree of accuracy, it was found that the ratio of the intensity of the central component to that of either of the outer components was certainly less than 1 2 1 From this it follows that the nuclear spin must be greater than $\frac{3}{2}$ (the ratios for spins of $\frac{1}{2}$ and $\frac{3}{2}$ being respectively 2 1 and 1 2 1)

The value of the intensity ratio of the two components of the line $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$ was determined with very great accuracy After allowing for all possible systematic errors it was found to be 1 21 1 From this it follows that the value of the nuclear spin is $\frac{9}{2}$, the intensity ratios corresponding to spins of $\frac{7}{2}$, $\frac{9}{2}$ and $\frac{11}{2}$ being respectively 1 29 1, 1 22 1 and 1 18 1 The values corresponding to spins of $\frac{7}{2}$ and $\frac{11}{2}$ are outside the probable limit of error

The intensity ratio of the components of the line $4S_{\frac{1}{2}}$–$5\,^2P_{\frac{3}{2}}$ was approximately measured The value was found to be 1 23 1, but the accuracy of this was not very high, an error of ± 0 06 being possible, the results of this measurement are however of some interest as they afford a certain confirmation of the value of the nuclear spin found from the intensities of the components of the line $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$

In a preliminary note Ritschl (1933) suggested that the nuclear spin of aluminium was $\frac{1}{2}$, the reason for this being an observed doublet structure in the lines 3057 and 3050 of the arc spectrum and the line $3S_0$–$3\,^3P_1$, 2669, of the spark spectrum However, a careful investigation of the structure of these lines showed that the doublet structure is only a partial resolution, the width of the components of the lines 3057 and 3050 being about twice as great as the normal width of a simple line under the conditions used The spark line 2669 was found to consist of one narrow line and another line of width twice as great This again clearly shows that the resolution into two components is incomplete, the observed structure being perfectly explained by assuming a triplet structure

The nuclear spin of $\frac{9}{2}$ found from the accurate measurement of the intensity ratio of the hyperfine structure components of the line $3\,^2P_{\frac{1}{2}}$-$3\,^2D_{\frac{3}{2}}$ is thus seen to be supported by the less accurate determinations of the intensity ratios of the components of the lines $4S_{\frac{1}{2}}$–$5\,^2P_{\frac{3}{2}}$ and $3\,^2P_{\frac{1}{2}}$–$4S_{\frac{1}{2}}$, and the apparent discrepancy of the doublet structures observed by Ritschl (which would indicate a spin value of $\frac{1}{2}$) is explained by incomplete resolution

Assuming the value $\frac{9}{2}$ for the mechanical moment of the nucleus, the magnetic moment can be calculated from the width of the splittings of the levels $4S_{\frac{1}{2}}$ and $3\,^2P_{\frac{1}{2}}$ Using the formulae given by Goudsmit, it is 4 1 nuclear magnetons according to the splitting of the level $4S_{\frac{1}{2}}$ and 3 6 nuclear magnetons according to the splitting of the level $3\,^2P_{\frac{1}{2}}$ The sign is positive, for in the lines $3\,^2P_{\frac{1}{2}}$–$3\,^2D_{\frac{3}{2}}$ and $4S_{\frac{1}{2}}$–$5\,^2P_{\frac{3}{2}}$ the components of longer wavelength are the stronger The discrepancy of 0 5 between the values of the magnetic moment calculated from the splittings of the levels $4S_{\frac{1}{2}}$ and $3\,^2P_{\frac{1}{2}}$ is much outside the limits of experimental error, however, Goudsmit states that the formula for P levels gives rather too low values of the magnetic moment with light elements

The authors take this opportunity of thanking Professor Plaskett for placing at their disposal his excellent photometer, and Professor Lindemann for his continued interest in the research, and also Queen's College and St John's College for the stipends granted to one of them

SUMMARY

The hyperfine structure of the resonance lines of aluminium was investigated by means of the absorption in an atomic beam of aluminium The line $3^2P_{\frac{1}{2}}$–$4S_{\frac{1}{2}}$ was found to possess three components at $-0\ 048$ $0\ 000$ and $+0\ 048$ cm $^{-1}$ of nearly equal intensity The line $3^2P_{\frac{1}{2}}$–$3^2D_{\frac{3}{2}}$ possessed two components of separation 0 066 cm $^{-1}$, the intensity ratio of which was measured to a high degree of accuracy, the value found was 1 21 1, the component of longer wave length being the stronger

From the structure of the line $3^2P_{\frac{1}{2}}$ $4S_{\frac{1}{2}}$ it follows that the levels $4S_{\frac{1}{2}}$ and $3^2P_{\frac{1}{2}}$ are both split into two levels, of separation 0 048 cm $^{-1}$ (the greater splitting of the line $3^2P_{\frac{1}{2}}$–$3^2D_{\frac{3}{2}}$ is due to a small unresolved, inverted structure of the level $3^2D_{\frac{3}{2}}$ which is probably caused by perturbation by the term $3s\,3p^2\,{}^2D_{\frac{3}{2}}$) The observed intensity ratio of the components of the line $3^2P_{\frac{1}{2}}$–$3^2D_{\frac{3}{2}}$ gives a value $\frac{9}{2}$ for the nuclear spin, and the splittings of the levels $4S_{\frac{1}{2}}$ and $3^2P_{\frac{1}{2}}$ give values 4 1 and 3 6 nuclear magnetons respectively for the magnetic moment of the nucleus according to the formulae of Goudsmit

The lines $4S_{\frac{1}{2}}$–$5^2P_{\frac{1}{2}}$ and $4S_{\frac{1}{2}}$-$5^2P_{\frac{3}{2}}$ were observed in emission and were found to be doublets, only just resolvable on account of their Doppler width, of separation about 0 05 cm $^{-1}$ and of intensity ratio (when corrected for overlapping) about 1 23 1, the component of longer wave-length being the stronger This structure is in complete agreement with the interpretation of that observed in the resonance lines

This result is in disagreement with the contents of a preliminary note by Ritschl in which, from an observed doublet structure of the arc lines 3057 and 3050 and the spark line 2669, he suggested a nuclear spin of $\frac{1}{2}$, an investigation of the structure of these lines showed that the observed doublet structures do not represent a complete resolution

REFERENCES

Jackson, D A 1930 *Proc Roy Soc* A, **128**, 508–22
Jackson, D A and Kuhn, H 1937*a* *Nature, Lond*, **140**, 110
— — 1937*b* *Proc Roy Soc* A **158**, 372–83
Ritschl, R 1933 *Nature, Lond*, **131**, 58–9

On Anomalous Vibrational Spectra

By M Blackman, *Trinity College, Cambridge*

(*Communicated by R H Fowler, F R S—Received* 4 *August* 1937)

In previous work on the vibrational spectrum of simple lattice models, it has been shown (Blackman 1937) that the spectrum becomes anomalous when special values are chosen for the force constants entering into the description of the model As an example we may take the case of a square lattice of lattice distance d containing one particle per cell, in which the force constants α (for particles at a distance α) and γ (for particles at a distance $d\sqrt{2}$) were used When γ/α tended to zero it could be shown that the spectrum changed from the two-dimensional to that of the linear chain A similar result holds in the corresponding three-dimensional case Although one would not expect any actual crystal to correspond to a limiting case (it would not be stable) it is conceivable that there should be crystals which approach the limiting case

In all cases which have hitherto been discussed, the anomaly is associated with the behaviour of the low-frequency end* of the spectrum This suggests that one could trace the effect to some property of the elastic continuum An examination of the lattice mentioned above, for which the elastic continuum has two elastic constants c_{11} and $c_{12} = c_{44}$, shows that the transition to the limiting case ($\gamma/\alpha = 0$) can be put in the form $c_{44}/c_{11} \to 0$

The investigation of the two- and three-dimensional continua shows that there are several relations among the elastic constants which lead to the velocity of elastic waves being zero in certain directions In all such cases the vibrational spectrum becomes anomalous The specific heat of those lattices for which the elastic constants satisfy the critical relations show of course a different character as well In all these cases the θ_D value is zero for very low temperatures It is hence clear that the study of the spectrum in such cases is of interest from the point of view of the specific heat

1—*Two-dimensional Theory*—We shall discuss the behaviour of a two-dimensional regular continuum, as this is similar to the three-dimensional case and has the additional advantage that the discussion of the limiting cases is simpler The equation of motion of the continuum can be solved on

---

* An anomaly can also be associated with the high frequency end of the spectrum, such cases are now being investigated

the assumption that we have wave motion and the velocity of the various waves can be found from the equation

$$\begin{vmatrix} -\rho c^2 + c_{11} l^2 + c_{44} m^2 & (c_{12}+c_{44})\, lm \\ (c_{12}+c_{44})\, lm & -\rho c^2 + c_{11} m^2 + c_{44} l^2 \end{vmatrix} = 0 \qquad (1)$$

Here l, m are the direction cosines of the wave, ρ is the density of the medium, c_{11}, c_{12}, c_{44} are its elastic constants, and c is the wave velocity

From the point of view of the spectrum it is convenient to look at this equation for the velocity in a slightly different manner We are interested primarily in the frequency ν and in the curves of constant frequency, so we put $\nu = c/\lambda$, $p = l/\rho\lambda$, $q = m/\rho\lambda$, where λ is the wave length, and obtain

$$\begin{vmatrix} -\nu^2 + c_{11} p^2 + c_{44} q^2 & (c_{12}+c_{44})\, pq \\ (c_{12}+c_{44})\, pq & -\nu^2 + c_{11} q^2 + c_{44} p^2 \end{vmatrix} = 0 \qquad (2)$$

The p, q are closely related to the phase differences between particles in the lattice for which (2) represents the limiting form of the equation of motion

Solving the above equation we obtain

$$2\nu^2 = (c_{11}+c_{44})(p^2+q^2) \pm [(p^2-q^2)^2(c_{11}-c_{44})^2 + 4(c_{12}+c_{44})^2 p^2 q^2]^{\frac{1}{2}} \qquad (3)$$

This can be put in a more convenient form by expressing p, q in polar co-ordinates, i e $p = R\cos\theta$, $q = R\sin\theta$, where $R = 1/\rho\lambda$, hence R is inversely proportional to the wave-length

$$R^2 = 2\nu^2/\{(c_{11}+c_{44}) \pm [(c_{11}-c_{44})^2\cos^2 2\theta + (c_{12}+c_{44})^2 \sin^2 2\theta]^{\frac{1}{2}}\} \qquad (4)$$

The two solutions correspond to the two frequency branches We can now plot the curves $R = R(\theta)$ for fixed values of the frequency and for different values of the elastic constants It will be seen that the maximum and minimum values of R occur at points for which $\theta = 0°$ or $\theta = 45°$ In these cases R has the following values

$$\left.\begin{aligned} R_1^2(0°) &= 2\nu^2/2c_{11} \qquad R_2^2(0°) = 2\nu^2/2c_{44}, \\ R_1^2(45°) &= 2\nu^2/(c_{11}+c_{12}+2c_{44}), \quad R_2^2(45°) = 2\nu^2/(c_{11}-c_{12}) \end{aligned}\right\} \qquad (5)$$

One notices immediately that the elastic constants must satisfy certain relations

$$\left.\begin{aligned} &(a) \quad c_{44} > 0, \\ &(b) \quad (c_{11}-c_{12}) > 0, \\ &(c) \quad c_{11} > 0, \\ &(d) \quad (c_{11}+c_{12}+2c_{44}) > 0 \end{aligned}\right\} \qquad (6)$$

These are simply the relations which are necessary in order to ensure that the frequencies should have real values (the condition for stability) When any one of the above combinations of elastic constants are equal to zero the value of R tends to infinity

We shall consider each of the above relations separately at first In the case $c_{44} \to 0$ the cross-sections $R = R(\theta)$ for a fixed value of ν are shown in fig 1 and in fig 2, for which $c_{44} = c_{12}$, $c_{12} = 0\ 79 c_{11}$ respectively The figures show how the critical curve is approached R becomes infinite along the axes in one branch only, whereas $R(45^\circ)$ is finite in both branches and depends on the value of c_{11} and c_{12}

One would not expect the value $R = \infty$ to be correct but would rather attribute it to a breakdown of the continuum theory It is interesting to compare the frequency equation with an exact one in this connexion The case $c_{12} = c_{44}$, $c_{11} \neq 0$ can be compared* with the square lattice mentioned in the introduction, the frequency equation for which is (Blackman 1935)

$$\begin{vmatrix} -m\omega^2 + 2\alpha(1-\cos\phi_1) & 4\gamma\sin\phi_1\sin\phi_2 \\ \quad +4\gamma(1-\cos\phi_1\cos\phi_2) & \\ 4\gamma\sin\phi_1\sin\phi_2 & -m\omega^2 + 2\alpha(1-\cos\phi_2) \\ & \quad +4\gamma(1-\cos\phi_1\cos\phi_2) \end{vmatrix} = 0, \tag{7}$$

where $\phi_1 = 2\pi dp/\lambda$, $\phi_2 = 2\pi dq/\lambda$ Expanding the periodic functions for large values of λ the frequency equation takes the same form as in (3) The expansion is, however, not always permissible and fails for example when the frequency is zero for a finite value of λ The cross-sections in the case $\gamma/\alpha = 0$ are straight lines in ϕ (or p, q) space Hence the limiting cross-section in fig 1 is incorrect though the others are correct For this particular lattice one knows that the whole spectrum becomes one-dimensional when $\gamma/\alpha = 0$, but this is a special feature and could not be predicted from the continuum equation What can be predicted is that the spectrum should have a one-dimensional character for low frequencies, i e the density of normal vibrations should be constant We know that the cross-section for one of the branches can be approximated by two straight lines parallel to the axes and that these lines have a finite length The area (for fixed ν) under the curve will hence be approximately represented by the sum of two rectangles less the area of the square common to the two rectangles One of the sides of the rectangles is a constant, the other is proportional to ν Hence the area can be written as $a\nu - b\nu^2$ In the case of the second branch the corre-

---

* This is only one of the many lattices which have the same continuum equation It is taken because it is a very convenient example

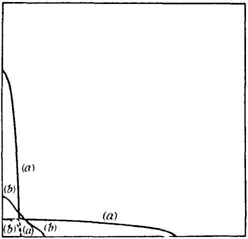


FIG 1—R θ diagram for the case $c_{44} \to 0$

(a) $c_{12} = c_{11}$, $c_{44} = c_{11}$,
(b) $c_{12} = 0.98 c_{11}$ $c_{44} = c_{11}$

The dotted curve is the second of the two frequency branches

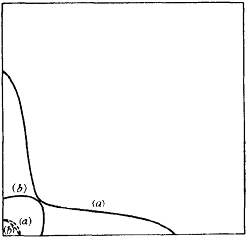


FIG 2—R θ diagram for the case $c_{44} \to 0$

(a) $c_{44}/c_{11} = 0.01$ $c_{12}/c_{11} = 0.79$
(b) $c_{44}/c_{11} = 0.20,$ $c_{12}/c_{11} = 0.79$

sponding area is $b\nu^2$ Hence the total area is $a\nu$ and the density $\rho(\nu)$ is given by the constant a In the general case in which $c_{12} \neq c_{44}$ the term in $b\nu^2$ in the area will not be cancelled and the density will be of the form $\rho(\nu) = a' - b'\nu$ *Hence we always have the one-dimensional density for small* ν

Taking the condition $c_{11} = c_{12}$ and assuming c_{44}/c_{11} of the order of unity, the curves of constant frequency take the form shown in fig 3 The anomaly now occurs for R_{45} It can be seen that this anomaly has certain properties in common with that of $c_{44} = 0$, so that one would be tempted to think that the density is again one-dimensional

It can again be shown by considering the simple lattice model, the frequency equation for which is given in (7), that the density is indeed exactly one-dimensional for small frequencies in the case $c_{12} = c_{11}$, $c_{44} = c_{11}$ This case can be obtained by putting $\alpha = 0$, $\gamma \neq 0$ The solutions of the frequency equation then become

$$\left.\begin{aligned} m\omega^2 &= 4\gamma(1-\cos\phi_1\cos\phi_2+\sin\phi_1\sin\phi_2), \\ m\omega^2 &= 4\gamma(1-\cos\phi_1\cos\phi_2-\sin\phi_1\sin\phi_2) \end{aligned}\right\} \qquad (8)$$

The cross-sections $\phi_1 = f(\nu\phi_2)$ are exactly the same as in fig 3, except that the curves are limited to a certain region in ϕ-space

But equation (8) can easily be transformed by putting

$$\phi_1+\phi_2 = \Phi_1, \quad \phi_1-\phi_2 = \Phi_2 \qquad (9)$$

into the form

$$\left.\begin{aligned} m\omega^2 &= 4\gamma(1-\cos\Phi_1), \\ m\omega^2 &= 4\gamma(1-\cos\Phi_2) \end{aligned}\right\} \qquad (10)$$

This is a one-dimensional frequency equation, and the whole spectrum is hence that of the one-dimensional lattice It follows that the density of normal vibrations is $\rho(\nu) = a + b\nu^2$ for small ν, the linear term being missing

Taking the third condition $c_{11}+c_{12}+2c_{44} = 0$ we can show that for $c_{44} = c_{11}$ exactly the same relations hold as in the case $c_{12} = c_{11} = c_{44}$ In each case $R^2(0°) = \nu^2/c_{11}$, ν^2/c_{44} and $R^2(45°) = \nu^2/8c_{11}$, ∞ Hence we have again a one-dimensional density distribution An analogous argument shows that the case $c_{11} = 0$ is related to the case $c_{44} = 0$

The remaining cases occur when the elastic constants are intermediate between those chosen for discussion above Two are of particular interest, they occur when two of the critical conditions are satisfied simultaneously

$$(a) \quad c_{11} = c_{12}, \quad c_{44} = 0,$$

$$(b) \quad c_{12}+c_{11}+2c_{44} = 0, \quad c_{44} = 0$$

As can be seen from (6) R becomes infinite for one of the branches in every direction

From the previous comparison between the results for particular lattices and the continuum one would suppose that this could only mean that the frequency is zero for the whole frequency branch It is rather interesting to see whether one can construct some lattice model which shows this property It should be noted that the first condition $c_{11} = c_{12}, c_{44} = 0$ fulfils the requirements for isotropy, so that one would look for some lattice which has a symmetrical grouping of particles One case which fulfils the conditions is a hexagonal network such as is found in graphite i e the particles are found only at the corners of the hexagon It can be shown that there are two particles in the unit cell, hence there are four frequency branches, two acoustical and two optical If the assumption is made that forces act only between neighbours, the equation for the frequency becomes

$$X^4 - X^2\alpha^2[12 + 2\cos\phi_1 + 2\cos\phi_1 + 2\cos(\phi_1 - \phi_2)] \\ + 9\alpha^4[3 + 2\cos\phi_1 + 2\cos\phi_2 + 2\cos(\phi_1 - \phi_2)] = 0, \qquad (11)$$

where $X = -m\omega^2 + 3\alpha$, α is a force constant and ϕ_1 and ϕ_2 the elementary phase differences

The solutions are

$$-m\omega^2 + 3\alpha = \pm 3\alpha,$$

$$-m\omega^2 + 3\alpha = \pm \alpha(3 + 2\cos\phi_1 + 2\cos\phi_2 + 2\cos(\phi_1 - \phi_2))^{\frac{1}{2}}$$

It will be seen that one of the roots is $m\omega^2 = 0$, i e all the frequencies for the branch are zero The other branches behave quite normally It can also easily be verified that the required relation holds between the elastic constants Since the condition of isotropy holds formally the vibrations should be purely transverse or purely longitudinal for long waves It is the transverse waves which behave anomalously here

The second of the above relations $c_{11} + c_{12} + 2c_{44} = 0$ $c_{44} = 0$ also* satisfies the condition for isotropy † there being two such conditions in the two-dimensional case in contrast to the three-dimensional case where there is only one Here it is the longitudinal waves which behave anomalously and give the frequency zero

In both the above cases the spectrum will have a strong maximum at the point $\nu = 0$ This can be represented as being an extreme case of the one dimensional spectrum for one of the branches $\rho(\nu) = a - b\nu$ in which a gradually tends to infinity in such a way that $\int \rho(\nu)\, d\nu$ is finite

---

* It is clear that the model mentioned above does not apply to this case

† The conditions are $A = (c_{12} + c_{44})/(c_{11} - c_{44}) = \pm 1$

This shows what happens in the intermediate cases The spectrum will in all cases be of the same form at the lower end, but the initial value of the density will rise enormously as the two critical points are approached

We have discussed above the different types of anomalous spectra in the case of a regular two-dimensional continuum The more complicated continua are also interesting but the effects though different in details can all be classed under those already mentioned, it is hence hardly worth while to extend the discussion

2—*Three dimensional theory*—As in the previous section we shall discuss the simplest types of elastic continuum and confine ourselves largely to the case of the regular continuum which is again characterized by three elastic constants c_{11}, c_{12}, c_{44} The equation of motion of the continuum can be reduced to a determinantal equation for the velocity of elastic waves which has the following form

$$\begin{vmatrix} -\rho c^2 + c_{11} l^2 + c_{44} m^2 + c_{44} n^2 & (c_{12}+c_{44})\, lm & (c_{12}+c_{44})\, ln \\ (c_{12}+c_{44})\, lm & -\rho c^2 + c_{11} m^2 + c_{44} n^2 + c_{44} l^2 & (c_{12}+c_{44})\, mn \\ (c_{12}+c_{44})\, ln & (c_{12}+c_{44})\, mn & -\rho c^2 + c_{11} n^2 + c_{44} l^2 + c_{44} m^2 \end{vmatrix} = 0, \tag{12}$$

where ρ is the density of the medium, l, m, n are the direction cosines of a wave in the medium, and c the velocity of propagation of the wave

As has already been explained we are interested in the frequencies rather than in the velocity, and we can change (12) into the following equation for the frequencies, putting $\nu = c/\lambda$, $p = l/\rho\lambda$, $q = m/\rho\lambda$, $r = n/\rho\lambda$

$$\begin{vmatrix} -\nu^2 + c_{11} p^2 + c_{44} q^2 + c_{44} r^2 & (c_{12}+c_{44})\, pq & (c_{12}+c_{44})\, pr \\ (c_{12}+c_{44})\, pq & -\nu^2 + c_{11} q^2 + c_{44} p^2 + c_{44} r^2 & (c_{12}+c_{44})\, qr \\ (c_{12}+c_{44})\, pr & (c_{12}+c_{44})\, qr & -\nu^2 + c_{11} r^2 + c_{44} p^2 + c_{44} q^2 \end{vmatrix} = 0, \tag{13}$$

We may then regard the equation as the limiting form, for low frequencies, of the equation for the frequency of normal vibrations of a lattice where p, q, r correspond to the phase difference between particles in the lattice In the p, q, r space we have then surfaces of constant frequency, there being three such surfaces corresponding to the three frequency branches It is necessary to investigate the points or cross-sections of the frequency surface for which the frequency becomes zero when the elastic constants are allowed to vary The general solution of (13) is rather complicated but it is sufficient to deal with special solutions as the anomalies are restricted to special directions in p, q, r space We shall consider the solutions for $p = q = r$ and those for $p = 0$, $q \neq 0$, $r \neq 0$, etc An examination of the general solution failed to show any other points worth investigation We shall express p, q, r in terms of polar co-ordinates $p = R\cos\theta\sin\phi$ $q = R\cos\theta\cos\phi$, $r = R\sin\theta$ In the case $r = 0$, $p \neq 0$, $q \neq 0$ the three solutions for R become

$$\left.\begin{aligned} R^2 &= 2\nu^2/\{(c_{11}+c_{44}) \pm [(c_{11}-c_{44})^2\cos^2 2\phi + (c_{12}+c_{44})^2\sin^2 2\phi]^{\frac{1}{2}}\}, \\ R^2 &= 2\nu^2/2c_{44} \end{aligned}\right\} \quad (14)$$

The first of these solutions is exactly the same as in the two dimensional case We shall write down the values in the special cases $\theta = 0°$ and $\theta = 45°$, as the R values take on extreme values for these cases

$$\left.\begin{aligned} &\theta = 0°, \quad R_1^2 = \nu^2/c_{14}, \quad R_2^2 = \nu^2/c_{44}, \quad R_3^2 = \nu^2/c_{11}, \\ &\theta = 45°, \quad R_1^2 = \nu^2/c_{44}, \quad R_2^2 = 2\nu^2/(c_{11}+c_{12}+2c_{44}), \quad R_3^2 = 2\nu^2/(c_{11}-c_{12}) \end{aligned}\right\} \quad (15)$$

The three solutions for $p = q = r$ ($\phi = 45°$, $\theta = \tan^{-1} 1/\sqrt{2}$) are

$$R_1^2 = 3\nu^2/(c_{11}-c_{12}+c_{44}), \quad R_2^2 = 3\nu^2/(c_{11}-c_{12}+c_{44}),$$
$$R_3^2 = 3\nu^2/(c_{11}+2c_{12}+4c_{44}) \quad (16)$$

It will be noted that two solutions fall together here, which means that two of the frequency surfaces touch for that direction in p, q, r space Furthermore, when $c_{11}-c_{12} = 2c_{44}$ all solutions fall into two groups for which $R_1^2 = \nu^2/c_{44}$, $R_2^2 = \nu^2/(c_{12}+2c_{44})$ respectively This is the isotropic case for which the surfaces of constant frequency are spheres, two of which fall together

From (15) and (16) one can construct a list of conditions which the elastic constants must fulfil

(*a*) $c_{44} > 0$,

*†(*b*) $(c_{11}-c_{12}) > 0$,

(*c*) $c_{11} > 0$,

(*d*) $(\frac{1}{2}c_{11}+c_{12}+2c_{44}) > 0$

---

* This condition has been given by Foersterling (1918) but its consequences were not investigated

† It has been suggested (Durand 1936) from a study of the variation of the elastic constants of NaCl, KCl MgO with temperature that the relation $c_{11} = c_{12}$ held at the melting point of these substances In view of the fact that the lattice is unstable when this condition holds this fact assumes a special interest, an examination of the experimental data shows however that the extrapolation is rather large, and further work would be necessary in order to establish this result

The condition $(c_{11}-c_{12}+c_{44})>0$ is automatically fulfilled when (*a*) and (*b*) hold, and similarly $c_{11}+c_{12}+2c_{44}>0$ if $c_{11}+2c_{12}+4c_{44}>0$

We can consider these conditions from the point of view of a diagram of c_{12}/c_{11} against c_{44}/c_{11} Then the relations (*a*), (*b*), (*d*) are straight lines in this diagram, which is useful in visualizing the connexion between the various conditions and also the range of values of the elastic constants which are forbidden This point will be considered in detail later

In the case where $c_{44}/c_{11}=0$, $(c_{11}-c_{12})/c_{11}>0$ an examination of (15) shows that the R values for one branch become infinite for both $\theta=0$ and $\theta=45°$, which means that they are infinite for the intermediate directions as well Hence the frequency is zero over a plane in the phase cube (Born 1923) for the corresponding lattice Since there are three planes which are equivalent this holds for three planes of the phase cube Furthermore, R becomes infinite for the second branch too in the direction of the space diagonal The third branch remains normal The question now arises as to what the density of normal vibrations becomes for the lattice when the above relation is fulfilled

If the position were such that the frequency is zero over only one plane and the surfaces of constant frequency are parallel planes whose distance is proportional to the frequency, then it is easy to see that the density of normal vibrations is a constant for small ν, i e the distribution is purely one dimensional If the frequency is zero along a line and the surfaces of constant frequency expand uniformly about this line as the frequency is increased, then the density is proportional to the frequency (for small frequencies), i e it is two dimensional

In the case $c_{44}/c_{11}=0$ the conditions are not quite so simple, but the fundamental point—that the frequency is zero over a plane and a line—still holds good One would hence expect a one-dimensional distribution for the first branch, a two-dimensional distribution in the second branch, while the third remains normal

These conclusions can be amplified by consideration of a special lattice which fulfils the condition with the added relation $c_{12}=c_{44}$ This is a simple cubic lattice of the Born-v Karmán (1912) type in which forces between neighbours only are considered It has been shown in previous work (Blackman 1937) that in this case the frequency equation has the form

$$\begin{vmatrix} -m\omega^2+2\alpha(1-\cos\phi_1) & 0 & 0 \\ 0 & -m\omega^2+2\alpha(1-\cos\phi_2) & 0 \\ 0 & 0 & -m\omega^2+2\alpha(1-\cos\phi_3) \end{vmatrix}=0 \tag{17}$$

For long waves this has the requisite form of the continuum equation with $\phi_1 = 2\pi dl/\lambda$, etc and $c_{11} = \alpha/d$ The frequency branches in this limiting case have been studied (Blackman 1937), and it has been shown that the density has the following form in the three cases

$$\text{(i)} \quad \rho_1(\nu) = a - b\nu - c\nu^2,$$
$$\text{(ii)} \quad \rho_2(\nu) = b\nu,$$
$$\text{(iii)} \quad \rho_3(\nu) = c\nu^2,$$

when the frequency is small This shows that the classification suggested is correct In the case where $c_{12} \neq c_{44}$ the coefficients b and c will not be the same in the three cases but the form will be unaltered

Taking the second condition $(c_{11} - c_{12})/c_{11} = 0$, $c_{44} \neq 0$ we see that only one of the $R(45^\circ)$ values is infinite, which means that R is infinite in three directions Going over to the phase cube of the equivalent lattice, this means in turn that the frequency is zero along three face diagonals of the cube This happens for only one of the branches, the other two behave normally By analogy with one of the cases discussed above we would assume in this case a density of the form $\rho(\nu) = b\nu - c\nu^2$, i e a two-dimensional distribution This case is also similar to the two dimensional case discussed in § 1, where the same relation between the elastic constants held The cross-sections of the frequency surface for which $p = 0$ would be very similar to those given in fig 3

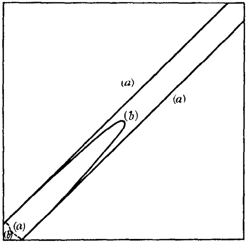


FIG 3—$R\ \theta$ diagram for the case $c_{12} \to c_{11}$

(a) $c_{12} = c_{11}$, $c_{44} = c_{11}$,
(b) $c_{12} = 0.98 c_{11}$, $c_{44} = c_{11}$

The difference between the two dotted curves is too small to be indicated Figs 1, 2, 3 are all drawn on the same scale for the same value of the frequency

The third case to consider is that where $(c_{11}+2c_{12}+4c_{44})/c_{11} = 0$, $c_{44} \neq 0$ which can occur only when c_{12} is negative Here it is one of the R values for $p = q = r$ which becomes infinite This means, in terms of the phase cube, that the frequency is zero along the space diagonal for one of the branches only It has hence a two-dimensional density

We have considered up to the present isolated points on each of the three lines in the c_{12}/c_{11}, c_{44}/c_{11} diagram, which denote the critical cases The lines cut in two points (i) $c_{12}/c_{11} = 1$, $c_{44}/c_{11} = 0$, (ii) $c_{44}/c_{11} = 0$, $c_{12}/c_{11} = -1/2$ Here the conditions become rather complicated as R becomes infinite (in all directions) for two branches in case (i) and for one branch in case (ii) The third branch is normal in (i), while in (ii) there is a line along which the frequency is zero as in all cases where $c_{44} = 0$

We have already discussed a similar case in two dimensions, so we know that the fact that R is infinite in all directions means that the frequency is zero for the whole of the phase cube Hence in (i)* two of the frequency branches have the frequency value zero By analogy with the two-dimensional case we would expect that it should be possible to find some loosely packed lattice which has the above property It is found that the diamond type will show these features if we consider the forces between neighbours only The analysis by Born (1914) shows this clearly, if we neglect in his work the forces which resist the change of angle

In case (ii) only one of the three branches disappears Here the transverse waves have the higher frequency (because c_{12} is negative)

The cases intermediate between the extremes will have intermediate properties We have already shown in two dimensions that one can reach the limiting case of a large heaping up of frequencies at $\nu = 0$ by letting the constant a in the expression $\rho(\nu) = a - b\nu^2$ become exceedingly large Similar considerations apply here

The surfaces of constant frequency and the cross sections for various frequencies show a similar variation in character (as the elastic constants are varied) to the two dimensional case, hence we can dispense with their discussion

3—One can study the critical conditions for the elastic constants in other systems as well as the regular system, but one can easily see that the types of anomalies in the spectrum will belong to one of the three classes as before, the large number of elastic constants make matters rather more complicated, so we shall discuss only one case—the hexagonal continuum—because there are a number of metals which belong to this system

---

* It is interesting to note that the equation of motion of the continuum goes over formally into that of the ideal liquid in case (i) The transverse waves "disappear" and only the longitudinal waves remain

The equation for the velocity of waves in the continuum with hexagonal symmetry is especially simple because the solutions are all symmetrical about the main axis We can hence study the determinantal equation for the frequency (or velocity) in any cross-section containing the axis, and the solutions will be the same for any other cross section We can therefore choose one of the direction cosines to be zero and can express the velocity of the waves in terms of the other two direction cosines (Zener 1936) The equations for the wave velocities are

$$\left.\begin{array}{ll} (a) & \tfrac{1}{2}(c_{11}-c_{12})m^2+c_{44}n^2-\rho c^2 = 0, \\ (b) & \begin{vmatrix} c_{11}m^2+c_{44}n^2-\rho c^2 & mn(c_{13}+c_{44}) \\ mn(c_{13}+c_{14}) & n^2c_{33}+m^2c_{44}-\rho c^2 \end{vmatrix} = 0 \end{array}\right\} \qquad (18)$$

Without going through the detailed calculations for the radius vector R we can see from (18) what conditions must be satisfied in order to obtain anomalies in the spectrum and what type of anomaly results We shall give in Table I the most important conditions denoting the anomalies by the numbers I, II and III I denotes a one dimensional, II a two dimensional frequency distribution while III indicates that the frequency is zero for a whole branch Finally N shows that the branch is normal

TABLE 1

Condition	Branch (*a*)	Branch (*b*)	Branch (*c*)
$c_{44} = 0$	I	I	II
$c_{12} - c_{11} = 0$	II	N	N
$c_{33} = 0,\ c_{11} = 0$	N	N	I
$c_{11} = c_{12},\ c_{44} = 0$	III	I	II
$c_{44} = 0,\ c_{33} = 0$	I	III	I
$c_{44} = 0,\ c_{11} = 0$	I	II	III

Besides the above there are other possibilities when the constant c_{13} has appropriate values These are of type II and will belong to one branch only

4—It is important to examine the above anomalies more closely from the point of view of the specific heat of the lattices Confining ourselves solely to the continuum region, we know that the θ_D value (which describes the specific heat) will tend to zero when the density of normal vibrations in the continuum region becomes very large Now the change from the normal density $\rho(\nu) = a\nu^2$ to an abnormal one $\rho(\nu) = a\nu$ or $\rho(\nu) = a$ is equivalent to the θ_D value tending to zero Hence we can study how the θ_D value tends to zero when the critical values of the elastic constants are reached

We shall confine ourselves to the case of the regular crystals as in § 2, and again make use of the diagram of c_{12}/c_{11} against c_{44}/c_{11}

The θ_D value can be determined from the density of normal vibrations and this in its turn from the elastic constants One can write this θ_D value in the form (Born 1923)

$$\theta_D = h\nu_D/k, \quad \nu_D = (\tfrac{3}{4}\pi)^{\frac{1}{3}}\,(\sigma/M)^{\frac{1}{3}}\,(c_{11}/\sigma)^{\frac{1}{2}} f(c_{12}/c_{11}, c_{44}/c_{11}), \tag{19}$$

where M is the mean atomic weight, σ the density and f is a dimensionless constant *The above relation for ν_D is practically identical with the Einstein* (1911) formula deduced from dimensional considerations In the formula given we can, however, assign a meaning to the constant $f(c_{44}/c_{11}, c_{12}/c_{11})$, it being a function of the volume in phase space enclosed by the surfaces of

TABLE II*

Substance	c_{11}	c_{12}/c_{11}	c_{44}/c_{11}	θ_D	M	σ	θ'
NaCl	4.77	0.276	0.276	302	29.25	2.18	1.68
KBr	3.327	0.174	0.185	179	59.5	2.49	1.57
KCl	3.75	0.053	0.175	227	37.25	1.98	1.53
KJ	2.67	0.16	0.152	132	83	3.08	1.48
NaBr	4.40	0.299	0.305	236	51.5	3.07	1.76
CaF_2	16.7	0.274	0.206	499	26.02	3.18	1.54
FeS_2	36.8	0.138	0.293	682	40	5.03	1.75
MgO	29.9	0.286	0.53	946	20.16	3.64	2.08
ZnS	9.43	0.603	0.46	326	38.72	4.06	1.71
Cu	18.6	0.73	0.44	342	63.57	8.94	1.59
Ag	12.0	0.75	0.37	212	107.9	10.5	1.55
Au	19.0	0.84	0.233	158	197.2	19.32	1.22
Al	10.8	0.59	0.29	398	27.1	2.6	1.57
Fe	23.6	0.595	0.481	461	55.84	7.86	1.80
Li	1.53	0.778	0.868	354	6.94	0.53	1.71
Na	0.972	0.853	0.596	143	23	0.97	1.43
K	0.447	0.861	0.582	77	39.1	0.80	1.33
W	51.2	0.40	0.298	384	184	19.12	1.67
Cd	16.1	—	—	305	112.4	7.0	1.48
Zn	12.1	—	—	189	65.38	8.7	1.31
Mg	5.18	—	—	360	24.32	1.75	1.75
Hg	3.6	—	—	68.5	200.6	14.19	1.14

* c_{11} is given in units of 10^{11} dynes/cm ^{2} The majority of the elastic constants are taken from Landolt Bornstein, "Physikalisch Chemische Tabellen" A few are from recent papers Those for Li, Na and K are theoretical elastic constants calculated by Fuchs (1936*b*) as are also the θ_D values (1936*a*) θ_D for NaBr, KBr and KJ as well as that for Mg were calculated by the author as none of these could be found in the literature

constant frequency when scale factors are removed We can write θ_D in the form*

$$\theta_D = \theta' c_{11}^{\frac{1}{2}} M^{-\frac{1}{2}} \sigma^{-\frac{1}{2}} h/k \tag{20}$$

We can now think of θ' as a surface in a space in which c_{12}/c_{11}, c_{44}/c_{11} are variables This surface will be bounded by the lines $c_{12}/c_{11} = 1$ $c_{44} = 0$, $2c_{12}/c_{11} + 4c_{44}/c_{11} = -1$, for which $\theta' = 0$ It is in general difficult to calculate θ' values but there are cases for which this can be done without much trouble, e g the isotropic case † Furthermore we can utilize the available data on θ_D values from elastic data A number of these have been calculated by different investigators for various substances One can reduce these by means of (20) to the θ' values It should be emphasized that the accuracy of the determination of the elastic constants does not enter into the discussion, we are concerned solely with the calculation of the θ_D value from a given set of data A comparison of the results obtained by different investigators (using different methods) shows a variation of several per cent so that the θ' values may be incorrect by 0.03 or so These θ' values have been plotted for other purposes in fig 4

The investigation of the θ' surface shows that there is initially a sharp rise from the value zero till a value of about unity is reached after which there is a slow increase The limiting value for large values of c_{12}/c_{11}, c_{44}/c_{11} depends very much on the direction from which it is approached, it may be zero (along the critical lines) a constant (in the isotropic case) or infinite

5—The distribution of crystals in fig 4 shows an interesting regularity namely the tendency of crystals of a particular type to group together The metals occupy the top section, the salts the lower left hand section If we devide the region $0 < c_{12}/c_{11} < 1$, $0 < c_{14}/c_{11} < 1$ into four blocks by means of the lines $c_{12}/c_{11} = 0.5$, $c_{44}/c_{11} = 0.5$ then we can put the polar crystals into the region I ($0 < c_{12}/c_{11} < 0.5, 0 < c_{44}/c_{11} < 0.5$), the face-centred metals into region II ($0.5 < c_{12}/c_{11} < 1, 0 < c_{44}/c_{11} < 0.5$), the body-centred metals into region III ($0.5 < c_{12}/c_{11} < 1, 0.5 < c_{44}/c_{11} < 1$), while region IV ($0.5 < c_{12}/c_{11} < 1$ $0 < c_{44}/c_{11} < 0.5$) is bare This classification applies actually only to the majority of cases, there are exceptions which encroach into adjoining regions

A more important point from our standpoint is the closeness of approach of the various crystals to the critical regions It would be difficult to obtain an exact criterion for this, but some idea might be obtained by considering the θ' value It gives some measure of the approach because the density of normal vibrations must become large when the elastic constants are near the critical values and hence the θ' value small, it being zero in the limiting case

---

* In this form θ' is not an invariant as it would be if we chose the compressibility instead of $1/c_{11}$, but as it is usual to express the elastic constants in the above form this does not appear to be a disadvantage

† Other cases are those where the condition for isotropy are nearly fulfilled, then one can use the formulae developed by Born and v Kármán (1913) Further, when $c_{12} = -c_{44}$ the equations of motion have a particularly simple form and the θ' value is easily found

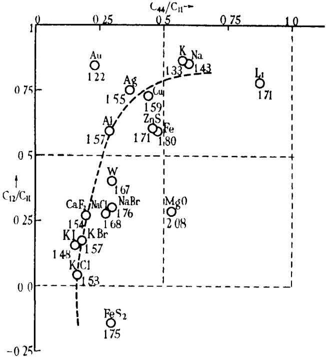


Fig 4—Diagram of c_{12}/c_{11} against c_{44}/c_{11} for the regular crystals Two of the critical lines are shown, $c_{12}/c_{11} = 1$, and $c_{44}/c_{11} = 0$ The third lies off the figure Each point is labelled with the corresponding θ' value The dotted line is a rough representation of part of the curve $\theta' = 1\ 57$

If we examine fig 4 from this point of view we see that the lower limit of θ' is independent of the particular structure, e g KJ (1 48), Au (1 22), K (1 33) The case of the alkali metals is interesting because it has been suggested that these are somewhat exceptional in their elastic behaviour and special features of the specific heat* were attributed to this cause This was obtained by studying the ratio $2c_{44}/(c_{11}-c_{12})$ (the "anisotropy ratio" as it is termed) which is large for the alkali metals (about 8) on account of the smallness of $c_{11}-c_{12}$ The face centred metals give a value for this ratio of about 4 If one were to adopt this ratio as another criterion of the approach to the critical region, it would have the value ∞ for $c_{11}=c_{12}$, and 0 when $c_{44}=0$ If one considers substances for which c_{44} is small one finds, e g , for KCl a value 0 4 It would seem difficult to compare such values with those for the alkali metals but none appear to be particularly close to the limiting values

---

* See Fuchs (1936*b*), also Mott and Jones (1936) The θ_D value of lithium shows a variation of about 30% It is puzzling to know why sodium and potassium should also be considered exceptional in this respect as the θ_D values are practically constant as far as they have been measured (Simon and Zeidler 1926) and this over a range of θ_D/T for which lithium shows a θ_D variation of 12%

It will be seen that the points lie in a comparatively narrow range of θ' values compared with the corresponding range of θ_D, these being 1 22–2 08 and 90–900 respectively, furthermore a large number are grouped about the value $\theta' = 1\ 57$ One can construct the curve $\theta' = 1\ 57$ over part of the range—it is the dotted curve in fig 4 One notices that it runs parallel to the axes over small regions for which θ' is independent of c_{12}/c_{11} or of c_{44}/c_{11}

Although none of the known crystals are near enough to the critical regions for any of them to be regarded as exceptional, it is quite likely that the vibrational spectrum of those which lie nearest the critical lines shows the approach to the critical region even if the specific heat does not The specific heat is controlled by large portions of the spectrum simultaneously, and not by any particular region except at low temperatures The approach would show itself in a flattening out of the spectrum in case I as the density must eventually become constant, in case II the density curve would tend to approach a line $\rho(\nu) = a\nu$ at the low frequency end As an example of a lattice approaching case I, we can take that of a Born-v Kármán lattice the spectrum for which has recently been worked out (Blackman 1937) Here we are rather nearer the critical line than for actual crystals ($c_{12}/c_{11} = 0\ 10$, $c_{44}/c_{11} = 0\ 10$) but the difference is not very great The flattening out of the spectrum is very apparent (as shown by fig 2 in the paper quoted) though we must allow for the fact that the lattice chosen is a very favourable one going over into the one dimensional case directly when the force constants are suitably chosen This feature of the spectrum is all the more interesting because the specific heat curve for the lattice shows little evidence of this one-dimensional character The initial value of θ_D is little lower than the high temperature value (see fig 4 in the paper quoted) for instance, nor are there any features of the θ_D-T curve which one could attribute to the one-dimensional character of the spectrum We see therefore that substances like KCl or sodium may show a one-dimensional or two-dimensional effect in the spectrum even if the specific heat is quite normal

6—In the previous section we have discussed the properties of substances with regular symmetry The only other case where some information is available is the hexagonal, the data for which has been given in Table II It will be seen that the θ' values are of the same order as the majority of the regular crystals It has also been shown in § 3 that the anomalies for the hexagonal case are of exactly the same type too One would therefore not expect anything new from a discussion of the θ' values for the hexagonal crystals This point is worth emphasis One might think that the fact that the forces in the direction of the main axis are sometimes different to those at right angles to it (as indicated by the smallness of c_{33} compared with c_{11} in the case of cadmium and zinc) shows that these substances cannot be put in the same class as the regular crystals, this is however erroneous The smallness of c_{33} indicates an approach to the condition III (see Table I) and is hence analogous to the approach to the relation $c_{11} = c_{12}$, $c_{44} = 0$ in the case of regular crystals

One would expect the specific heat curves for the hexagonal crystals to be of the same type as those of the other metals Actually zinc and cadmium show a rather large variation in θ_D, but this is not larger than that of lithium, while magnesium has a θ_D-T curve rising with temperature which can be matched by the case of tungsten

I should like to express my thanks to Professor R H Fowler and to Dr R Peierls for their interest I am indebted to the Department of Scientific and Industrial Research for a 'Senior Research Award'' which has made the above work possible

SUMMARY

The equation of motion of an elastic continuum is investigated, and it is shown that the continuum becomes unstable when the elastic constants obey certain relations The vibrational spectrum of lattices for which these relations hold is shown to undergo a characteristic change, the density of normal vibrations going over into the one-dimensional or into the two-dimensional form The significance of this is considered from the point of view of the specific heat and the vibrational spectrum of actual crystals, it is found that no actual crystals are near enough to the critical regions for the effect of the approach to be apparent in the specific heat, but in some cases this should be noticeable in the vibrational spectrum

REFERENCES

Blackman, M 1935 *Proc Roy Soc* A, **148** 384
— 1937 *Proc Roy Soc* A, **159**, 416
Born, M 1914 *Ann Phys Lpz*, **44** 605
— 1923 "Atomtheorie des festen Zustandes"
Born, M and von Kármán, Th 1912 *Phys Z* **13** 297
— — 1913 *Phys Z* **14** 15
Durand M 1936 *Phys Rev* **50**, 453
Foersterling, K 1918 *Z Phys* **8**, 25
Fuchs, K 1936*a* *Proc Roy Soc* A **153**, 622
— 1936*b* *Proc Roy Soc* A, **157**, 444
Mott, N F and Jones, H 1936 "Theoretical Properties of Metals and Alloys"
Simon, F and Zeidler, W 1926 *Z phys Chem* B, **123**, 383
Zener, C 1936 *Phys Rev* **49** 16

Crystal Growth from Solutions

By W F Berg, Ph D (Berlin), Ph D (Manchester)

(*Communicated by W L Bragg, F R S —Received 21 August* 1937)

[Plate 2]

INTRODUCTION

The purpose of the present paper is to provide an experimental contribution towards the knowledge of the mechanism of crystal growth from solutions

Certain details about the mechanism of crystal growth have been studied in the case of one component systems, i e crystals growing from the melt or from the vapour Very little is known about crystal growth from solutions Yet, although the one component system is the simpler one, special interest is attached to the two-component system, because of its technical applications In chemical industry crystals are grown from solutions on a big scale A closer knowledge of the mechanism of crystal growth would enable one to apply the conditions, e g temperature or degree of supersaturation, most favourable for the process in question This is, as yet, done purely empirically

The problem covers a wide field, and an attempt to answer what was thought to be one of the first questions to be asked will be described below

This is the question of the connexion between the rate of growth and the concentration of the solution in contact with the growing crystal face It will be seen that this question cannot be answered yet, mainly because impurities play too big a part in determining the rate of growth But the method applied seems to reveal certain facts about the mechanism of crystal growth from solutions

The concentration in actual contact with a dissolving or growing crystal has, since Nernst (1904, see also Brummer 1904), always been thought of as being the saturation concentration The rate of growth was then supposed to be proportional to the degree of supersaturation Several investigators have tried to prove these ideas experimentally Their method generally consisted of putting a known amount of seed crystals into a solution of known supersaturation and of separating the crystals from the solution after a certain time and of analysing it The solution was kept in rapid motion, and it was hoped to get solution of the measured concentration right up to the crystal faces, so that the concentration gradient became very big and diffusion processes could be neglected It seems very difficult to judge whether this has been achieved or not It is still possible that in these experiments the rate of growth was mainly governed by the rate of diffusion through a thin stagnant film of solution adhering to the crystal surface The results of the different papers are not in agreement The rate of dissolution was found by Noyes and Whitney (1897) to be proportional to the degree of unsaturation Marc (1905, 1908*a*) found the rate of growth to be proportional to the square of the supersaturation His results were not confirmed completely by later investigators (Le Blanc and Schmandt 1911, 1913, see also Jenkins (1925) and Roller (1932)) Experiments on the influence of dyes on crystal growth led Marc (1908*b*, 1910, 1911*a*) to assume the existence of an adsorbed layer on the surface of a crystal in contact with its solution This layer forms an intermediate state between the liquid and the solid Molecules go into this layer and stay there until the conditions become favourable for them to freeze into the lattice Dye molecules are adsorbed by the crystal, thereby changing the rate of growth of the particular faces adsorbing them and thus changing the habit of the crystal

Experiments by Volmer (1932) performed, however, on crystals growing from the vapour or from the melt, demonstrated the existence of such an adsorbed layer and its importance for the mechanism of crystal growth It was shown that the molecules in the adsorbed layer have a much higher mobility than in the melt

It appears that many questions concerning crystal growth need clearing

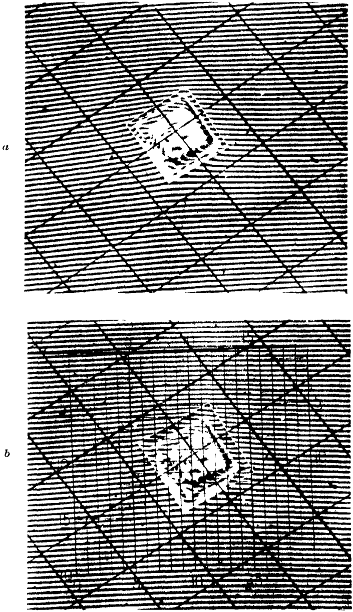


Fig 2—Shows two out of a series of photographs from a growing crystal The big graticule was photographed at the same time as the crystals the length of the sides of the squares corresponding to 1 mm The small graticule was attached afterwards for measuring purposes *a* is the photograph from which the diagram fig 6 was obtained The photographs are fixed in the same position as the diagram the lines are running horizontally The photographs are negatives

(Facing p 80)

up, especially in view of the existing discrepancies Complete changes in the method of attack and experiments on single crystals seem to be called for, since diffusion effects have to be eliminated or taken into account

METHOD

A method in which the concentration can be measured everywhere in the neighbourhood of a growing or dissolving crystal has been suggested to the writer by Mr T R Scott, of I C I (Alkali), Ltd, Winnington In this the problem is first reduced to a two-dimensional one, by placing a thin prismatic crystal of rectangular cross section between two horizontal glass plates, which stop growth in the third dimension Secondly, convection currents are eliminated by doing the experiments on a microscopic scale Thus the transport of molecules is purely a two dimensional diffusion process Thirdly, the two glass plates, forming a slight wedge, are half-silvered and illuminated with semi-parallel monochromatic light, thus producing a system of interference fringes These fringes are straight if the wedge is optically perfect and if the refractive index of the matter in the wedge is constant They are curved if the refractive index changes from place to place and thus provide a simple means of measuring the refractive index The refractive index will vary if a growing crystal takes molecules from the surrounding liquid Since top and bottom of the crystal touch the glass plates, symmetry considerations show that the diffusion process will be identical in layers parallel to the plates Thus the concentration distribution can be measured right up to the crystal face, if the refractive index is known in terms of the concentration Measurements by Miers and Isaac (1906) give the necessary figures for sodium chlorate the refractive index varies linearly with the concentration Sodium chlorate was chosen for the following experiments, because it crystallizes in rectangular prisms, and is a cubic crystal Supersaturated solutions of it are also easily obtained

APPARATUS

The apparatus consisted essentially of a photomicrographic outfit There are, however, details which seem worth mentioning (fig 1)

A "Patna" microscope by Watson was used, which had a specially made short and wide body tube, to avoid cutting down the field of view Microscope, camera, and source of light were all mounted on the cylindrical bed of a 4 in Drummond lathe This lathe proved extremely convenient for optical purposes The lens used for photographing was a *F* 3 5 Ross Xpress of 5 cm focal length No eyepiece was used for photographing

It appeared desirable to photograph a system of fiducial marks at the same time as the growing crystal and its surrounding fringe system This would facilitate comparison of different plates and would also give at once the magnification For this purpose a luminous graticule was used, which was obtained by contact printing from an eyepiece micrometer graticule

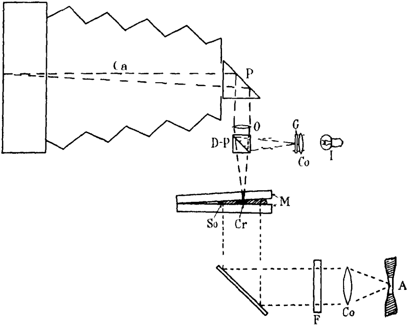


Fig 1—Shows the apparatus for photographing the growing crystals and the surrounding interference system *A* is the mercury arc, *Co* a condenser, *F* a filter, *M* are the two half silvered mirrors with the crystal *Cr* and the solution *So* between them *L* is a pea lamp, *Co* another condenser, *G* the graticule with light lines on a dark ground *D*–*P* is the double prism acting as transparent mirror *O* the object glass, *P* a right-angle prism and *Ca* the camera

A slightly silvered mirror was fixed underneath the objective so that the object and the graticule could be seen at the same time The mirror consisted of two right-angle prisms, one of which was slightly silvered, stuck together with canada balsam The graticule was fixed in front of a small condenser and a pea lamp, and the whole system could be moved towards the mirror to focus it at the same time as the object

The illumination for the interference phenomenon was provided by a small mercury arc taking 0 25–0 4 amp The light was made monochromatic by the use of filters The actual arc burned in a capillary which was so narrow that a symmetrical achromatic condenser made a beam of light sufficiently parallel to produce sharp interference fringes

Photographs were taken on Ilford Hypersensitive Panchromatic plates, because of their high speed for the mercury green line $\lambda = 5461$ A For visual observations when adjusting the apparatus this line was the brightest and most convenient one Exposure times ranged from 2 to 30 sec , applying a magnification of about 10 linear

The good quality of the interferometer plates and their silvering was essential to success The plates were 1 by 3 in and $\frac{1}{4}$ in thick, worked by Bellingham and Stanley "correct to $\frac{1}{4}$ wave-length" The silvering was done by cathode sputtering (Tolansky and Lee 1936)

EXPERIMENTS

The actual experiment consisted in placing one seed crystal in a drop of supersaturated solution between the two optical flats and taking a series of photographs of the crystal and the interference system at known times

Flat rectangular seed crystals about 1 mm square and 0 1 mm thick were obtained by slowly evaporating a drop of sodium chlorate solution on a piece of glass plate The degree of supersaturation of the solution was obtained by weighing The crystalline material was supplied by the British Drug Houses as having the following analysis

Chloride	0 02 %
Sulphate	Nil
Calcium	Nil
As_2O_3	0 8 part per million
Heavy metals, etc	Extremely faint trace

The main difficulty in the experiments was to avoid the formation of unwanted nuclei To stop this several precautions were applied The material was recrystallized twice, always discarding the first and the last portion It was hoped that physical impurities would thus be removed All the water used was distilled water and was filtered through filter crucibles before being used Ordinary filter paper caused trouble by giving off fluff

Despite all these precautions it was extremely difficult to conduct an experiment so that the seed crystal was the only one present After many trials it was found that these unwanted crystals came from the surface of the seed crystal It appears that when the solution round the growing crystal dries up on preparation of the seed crystal, the material of the last drop does not freeze into the structure of the crystal, but forms many very small crystals on top One might perhaps get over this by making evaporation extremely slow Another procedure was applied here It consisted in washing the seed crystal just before it was used in a drop of very slightly unsaturated solution The wet crystal was then transferred to one of the mirrors, the unsaturated solution soaked away with a piece of filter paper, the supersaturated solution applied with a small pipette and the whole covered with the second mirror The utmost speed was necessary, otherwise more nuclei were formed

The obvious disadvantage of getting rid of unwanted nuclei in this way is that the degree of supersaturation of the liquid surrounding the crystal at the start of the experiment is unknown, since it is not the same as the supersaturation of the solution one applies Some of the unsaturated solution may still stick to the crystal and mix with the supersaturated liquid It will be seen below how this difficulty was overcome by measuring the refractive index far away from the crystal

Although few precautions were taken the room temperature during the experiments only changed a few tenths of a degree centigrade The room temperature was measured each time a photograph was taken

When the seed crystal with its surrounding supersaturated solution had been placed between the two mirrors, the whole arrangement was transferred to the microscope stage and observed, using monochromatic light The illumination was adjusted till the fringes were sufficiently sharp, and when it was ascertained that no unwanted crystals were present the eyepiece used for visual observations was taken out and fringe system and graticule focused on the screen of the camera The first photograph was usually taken 3–4 min after the experiment was begun This and the last photograph always included the boundary of the drop Fringes appeared in the solution and in the neighbouring air, and by counting them over a distance the ratio of the refractive index of air and of the solution could be obtained This is a simple means of measuring the absolute value of the concentration far away from the crystal, which had to be done, because owing to the method used the degree of supersaturation was not known at the start and would most certainly change during the experiment The other photographs, six or eight in the course of 1 or 2 hr , were taken at different measured intervals with the crystal in the centre of the photographs

The thickness of the wedge was found by measuring the thickness of the crystal after the experiment This was done by removing the top mirror, drying the crystal superficially and focusing on the top and bottom of the crystal, using high magnification and the fine adjustment of the microscope By comparing the photographs taken at different times it could be ascertained whether the thickness of the wedge had changed, i e whether the crystal had grown in the third dimension, and if so by how much The fringes will move twice the number of fringe distances as the thickness changes in wave-lengths, thus providing a rather sensitive test No appreciable change in thickness of the wedge was ever found

EVALUATION OF THE PHOTOGRAPHS AND SOURCES OF ERROR

Fig 2 (plate 2) shows typical examples of the photographs obtained The curvature of the fringes as they approach the crystal is clearly visible

The shift of the fringes is now to be evaluated in terms of changes in refractive index and concentration The optical path difference between a ray going straight through the wedge and one reflected once at the upper and once at the lower mirror is $2\mu_0 d$ and this has to equal $N\lambda$ for a bright fringe

$$2\mu_0 d = N\lambda \tag{1}$$

If the fringe system is displaced by n fringe distances because μ_0 has changed to μ_1, we have

$$2\mu_1 d = (N + n)\,\lambda \tag{2}$$

and thus

$$\mu_0 - \mu_1 = -n\lambda/2d, \tag{3}$$

where μ_0 = the refractive of the supersaturated solution at a distance from the crystal,
μ_1 = the changed refractive index,
d = the thickness of the wedge,
N = an integer,
λ = the wave-length of the light

λ is known and d is obtained experimentally Thus, to obtain the concentration at any point on a fringe, it is only necessary to determine the displacement of the fringe at that point from the straight line formed when no crystal is present If only relative concentrations are wanted, as, for example, if only the concentration gradient distribution is interesting, it is sufficient to measure and discuss the fringe displacement only without knowing the absolute value of the refractive index and of the concentration

The concentration gradient can also be obtained more directly than by measuring the concentration distribution One has only to measure the angle between the tangent to a deviated fringe at the required point and the original straight line, and the distance of the fringe from its neighbours Since this method will be applied most frequently close to a crystal boundary it is more convenient to measure the spacing of the fringes along the crystal surface and the angles between the crystal and the fringes (see fig 3) Let

x_N = x-coordinate of the Nth fringe,
a = a constant,
b = spacing of the fringes in the x-direction (undeviated),
γ = tan of the angle α between the z-axis and the undeviated fringes

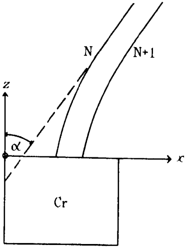


FIG 3—Illustrates the calculation of the normal concentration gradient from the curvature and the distance of the fringes The crystal *Cr* and two fringes *N* and *N* + 1 are drawn, whose distance (undeviated) measured parallel to the *x* axis is *b* If the fringes were not deviated they would make the angle *α* with the *z* axis

We can then derive a formula for the component of the concentration normal to the crystal face $\partial c/\partial z$ This component is of greatest interest, since it will be shown below that it is proportional to the current of molecules going into the crystal We find

$$\frac{\partial c}{\partial z} = \frac{\gamma}{\alpha} - \frac{1}{a}\left(\frac{dx}{dz}\right)_{z=0} \frac{b}{(x_N - x_{N-1})_{z=0}} \tag{4}$$

Thus we have to measure the inclination of the fringes with respect to the crystal face $\left(\frac{dx}{dz}\right)_{z=0}$ and the fringe distance $x_N - x_{N-1}$, and the distance of the undeviated fringes b which can be obtained far away from the crystal, where the concentration is still the original one The distribution of $\partial c/\partial z$ is then known apart from an unknown constant a This method will not be used for quantitative results, since the accurate measurement of the angle is rather difficult, but it enables one to judge at a glance how the concentration gradients are distributed

The measuring up of the diagrams was done in two ways At first the diagram was brought into contact with a graticule of twenty one horizontal "lines" and twenty-one vertical "columns", the lines being made to run parallel to the undeviated fringes far away from the crystal (see fig 2) The positions of the fringes with respect to this graticule were measured by means of a microscope and an eyepiece micrometer scale Later on more accurate measurements were done by means of a two-dimensional comparator The eyepiece of the microscope of the comparator contained two parallel lines of variable distance, and in measuring the fringe was "caught" between those two lines This procedure increased the accuracy appreciably, but produced no qualitative changes from the earlier method in the results given below

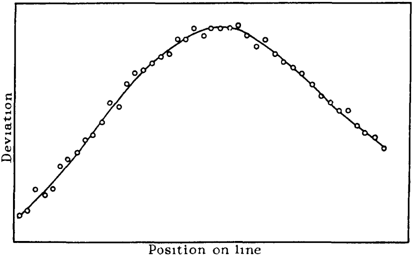


FIG 4—Gives an example of the concentration distribution along a "line" Each point represents the position of a fringe against which is plotted the deviation of this fringe from straight line configuration

The deviations of the fringes, i e the difference between their position and the straight-line configuration as determined far away from the crystal, where the fringes were not yet shifted, was plotted against the position of the fringes along the lines Fig 4 shows a typical example The positions of the fringes were also measured along the crystal surface (fig 5) Obviously the latter values are of special interest

Relative concentration diagrams were obtained from these curves in the following way Those coordinates on a certain line were found, where the deviation had certain arbitrarily fixed values, the difference between these values being constant Thus, for example, the coordinates may be found, where the deviation is 1, 2, 3, etc , fringe distances The coordinates obtained in that way were plotted on a graph, which was a reproduction of the graticule with its twenty-one lines and twenty-one columns, and

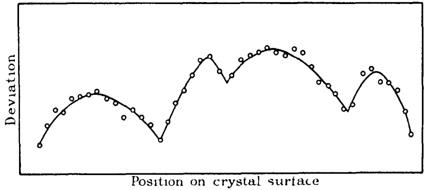


Fig 5—Shows the concentration distribution in arbitrary units (fringe deviations of the solution in contact with the four faces of a crystal

with the crystal drawn in in its proper position Thus a system of points of constant concentration was obtained and lines of constant concentration could be drawn through these points There were, however, places in the diagram where the points were too far apart to draw curves through them with any certainty, namely, where the lines of constant concentration run parallel to the lines of the graticule To fill up such gaps the concentration distributions along the columns were wanted These were obtained by taking the concentration values for a certain column from the curves for the various lines

In the final concentration diagram the different sorts of points were marked in a different way, because they carry a different weight The points on columns (●) are obtained by a twice-repeated smoothing out process (drawing a curve through points) and therefore carry the biggest weight, assuming that there are no systematic errors Points on lines are subject to only one such process (○) Points on the crystal surface ( ) carry even less weight, because each of them was measured separately and is therefore subject to the error attached to a single measurement

The accuracy of these concentration diagrams is not as high as is desirable This is mainly due to the fact that the flats used were only accurate to $\frac{1}{4}$ of a wave-length Much trouble was caused by this fact, particularly in one case, where an inexplicable distortion of the fringe system was finally traced back to the existence of a pit in one of the flats As soon as the crystal had grown across the pit and the fringes were in a different part of the wedge the distortion disappeared It would have been quite impossible to obtain any results at all if the accuracy of the flats over the small areas used had not been actually much higher The necessity of using the highest grade flats obtainable for any future experiments is emphasized by the fact that the accuracy of measuring the position of a fringe may under favourable conditions be as high as $\frac{1}{50}$ of a fringe distance, i e changes in the wedge thickness of $\frac{1}{100}$ of a wave-length would be detectable

Systematic errors may arise from the heat of crystallization The boundary of a growing crystal is a source of heat Any local change in temperature would distort our fringe system It can be shown by simple calculations, however, that in our case the rise in temperature on the crystal boundary will be between 10^{-4} and 10^{-2}° C only

Other errors may arise from the heat from the illuminant being absorbed by the crystal and its surroundings, resulting in a distortion of the fringes As a check experiment a crystal was put into saturated solution, and it was found that the fringes showed hardly any observable curvature What little shift there was was explained by evaporation from the boundaries of the drop, and the shift was the same all over the crystal, a point which will be important later on

It seems justified to treat our concentration diagrams as if there were no systematic errors

DISCUSSION OF THE RESULTS

Fig 6 shows a typical example of the concentration diagrams obtained

We are now in a position to discuss our original question of how the rate of growth depends on the degree of supersaturation in contact with the crystal surface It becomes clear at once by looking at the diagram fig 6 and at fig 5 that the question put forward in this simple way has no meaning, because the concentration in contact with the crystal varies from place to place This has been found to be always the case and is of great importance in view of the older suggestion that a layer of definite concentration surrounds the whole crystal (Nernst 1904) The concentration was always found to be greater near the corners of the crystal than at the centres of its faces Thus it seems a more appropriate question to ask how the rate of growth depends on the maximum, or minimum, or average, concentration in contact with the crystal face But even such a question is, for experimental reasons, extremely difficult to answer An example may illustrate this

It has occasionally been found that two opposite, i e in our case

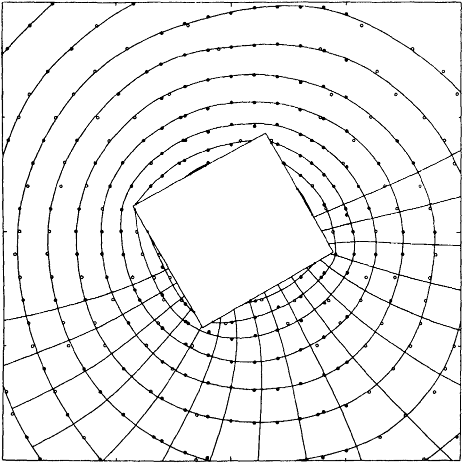


FIG 6—Shows a typical concentration distribution round a growing crystal The circles are points on "lines" (○), the large dots points on "columns" (●), the small dots ( ) points measured on the crystal surface "Lines" run horizontally, "columns" vertically in the diagram Lines of flow are drawn in the lower part of the diagram Note that the spacing of the lines of flow and of the lines of constant concentration is practically the same near the corners as at the centres of the faces The photograph from which this diagram was obtained is shown in fig 2*a*

geometrically and crystallographically equal, faces of the same crystal, which were in contact with solution of the same degree of supersaturation at the start, grew at different rates For this very reason the slower growing face was after a time in contact with more highly supersaturated solution than the other one Small crystallization velocity means higher concentration on the crystal surface under otherwise similar conditions The explanation is that the crystallization velocity is influenced by minute traces of impurities Hence consistent results have not yet been obtained and it will not be easy to obtain them

More interesting results were obtained when the diffusion gradients at different parts of the same crystal face were compared In this connexion it was interesting to note that in our experiments stationary conditions were reached almost at once Only during the first moments after the experiment had been started was there any appreciable movement of the fringes The graticule photographed on each plate proved very useful here Two successive photographs could be laid on top of each other, the graticules registering, and the shift of the fringes observed If such a shift had occurred at all it was slight compared with the deviation of the fringes from the straight-line configuration Exceptions were cases where the boundary of the drop of supersaturated solution was quite unsymmetrical with respect to the seed crystal Then the solution on the one side lost its supersaturation much quicker than on the other side, and the fringes moved accordingly It is self evident that there will be a slight continuous shifting of the fringes all the time, since the dimensions of the drop are finite and the supply of molecules is therefore limited The fact that stationary conditions prevail makes it quite simple to work out from our concentration diagrams the currents of molecules diffusing towards the crystal A two-dimensional diffusion equation

$$\frac{\partial c}{\partial t} = -D\nabla^2 c \tag{5}$$

holds for the general case, but this is reduced to

$$\nabla^2 c = 0 \tag{6}$$

for our stationary conditions The flow of molecules j is given by the equation

$$j = -D \operatorname{grad} c \tag{7}$$

The diffusion coefficient for sodium chlorate has not yet been determined, but it can in principle be found from any set of concentration diagrams

For the results described below the knowledge of the diffusion coefficient is not necessary

The fact that stationary conditions can be assumed provides us with an easy means of checking the validity of the diffusion equation in our case, ensuring that no systematic errors fake our results According to equation (6) the lines of flow and the lines of equal concentration are perpendicular to each other, and the spacing of the lines of flow is proportional to the spacing of the lines of equal concentration The density of the lines of flow gives the flow per unit area

These relations can, of course, only be verified graphically, since the equation cannot be solved analytically except for a few special boundary conditions To prove the relations lines of flow were constructed as indicated above They were found to run perpendicular to the lines of equal concentration fairly satisfactorily if their spacing was made proportional to the spacing of the lines of equal concentration The lines of flow in fig 6 were obtained in this way The accuracy with which this law was fulfilled gave little reason to doubt the validity of the diffusion equation for our case or of the concentration measured in the way described Any deviations from the law were not of a systematic nature Sometimes the lines of flow were a little too close, sometimes a little too far apart, which was explained by imperfections of the wedge It seems thus justified to work out the flow of molecules from the concentration diagrams by applying the diffusion equation But this result means more It should be possible to construct the whole concentration diagram from the knowledge of the position of two neighbouring lines of constant concentration Although one would never push it thus far this conclusion is certainly a great help in evaluating concentration diagrams, if due to local imperfections the diffusion equation does not seem to hold over the whole diagram

We can now proceed to work out the current distribution over the crystal faces A crystal face has generally the same rate of growth all over, as is evident from the fact that it remains plane, and only such faces were considered This supposition is easily verified by inspecting the different photographs taken It will be interesting to see whether the current of molecules flowing towards a crystal face is the same all over or not If the flow of molecules per unit area of the crystal is the same at every point one would say that the molecules are taken out of the solution by the growing crystal just where they are wanted to produce plane crystal faces If the flow is found to differ from place to place there must be another mechanism to make an even speed of growing possible

The flow per square centimetre or, in our two-dimensional case, per unit length of the crystal is the quantity to be compared, this is proportional to the component of the concentration gradient normal to the crystal face The concentration gradient is in turn inversely proportional to the spacing of the lines of constant concentration

The result of such a comparison is, as is easily seen from the concentration diagram reproduced, that the component of the concentration gradient normal to the face does differ from place to place We see that this concentration gradient is smallest near the corners and rises to a maximum at the centre of a face The absolute value of the gradient itself does not differ much from place to place, at least not over one face, so that one is tempted to assume as boundary condition that the absolute gradient is constant Higher accuracy will be necessary to confirm this Since the lines of constant concentration cut the crystal at about 45° at the corners, the normal component of the gradient, i e the flow of molecules per unit area of the crystal surface, is deficient at the corners

FIG 7—Shows a much enlarged portion of a photograph of a crystal and its surrounding fringes to demonstrate the difference in angle between the fringes and the crystal surface near the corner and at the middle of the face

The suspicion may arise that a systematic error in the evaluation has caused these results But, apart from our discussion on possible errors, as has been shown above, there is the possibility of estimating the normal component of the concentration gradient from the curvature and the distances of the fringes, and if the results are true we should notice that the curvature and the spacing of the fringes is smaller near the corners of the crystal than at the centres of its faces This is indeed the case as is demonstrated in fig 7

A new mechanism has to be assumed which enables the crystal to produce plane faces, although the current of molecules arriving at the corners is smaller than at the centres of the faces The only possibility seems to assume a lateral surface flow of molecules from the centre of a face towards the corners This current must flow in a very thin layer, because otherwise it would cause optical effects and would show up in our photographs by a shift of the fringes This result has been obtained on all crystals so far measured The decrease in the flow of molecules near the corners was found to be of the order of 25–50 % of the flow near the centres That means that this percentage of the molecules going into the corner of a crystal must flow part of their way in the thin layer

We are thus led to assume the existence on crystals in contact with their solutions of a layer which constitutes a transition between the solid and the liquid state We may call it an adsorbed layer in order to signify that it is very thin without implying that it is a monomolecular layer as seems to be quite often the case with adsorbed layers This adsorbed layer seems to be all-important for the mechanism of crystal growth Its existence has already been postulated by Marc (1908*b*, 1910, 1911*a*) and has been demonstrated by Volmer (1932) in the case of metal crystals growing from the melt or from the vapour

There seem to be certain discrepancies in these results One would be inclined to assume that the exchange of molecules between the solution and the adsorbed layer would be governed by laws of probability the probability that a certain number of molecules goes over in a certain direction per unit time should be proportional to the number of collisions, i e to the concentration of the molecules, all other factors being equal But here we find that *more* molecules go into the adsorbed layer at the centre of a face, where the concentration of the solution is *smaller* than at the corners, and where the concentration in the adsorbed layer must be greater than at the corners to produce the requisite lateral flow If our interpretation of the concentration diagrams is correct the only solution out of this difficulty is to assume that there exists a hitherto unknown effect which makes the molecules in the solution and in the adsorbed layer react differently on the corners and at the centres of the faces Lacking further experimental material no speculations will be made here as to the nature of this effect

The writer's thanks are due to Imperial Chemical Industries, Ltd , London, for a grant which made this research possible, and to Mr T C Swallow and Mr T R Scott, of I C I (Alkali), Ltd , Winnington, for suggesting the problem He also wishes to express his gratitude to Professor W L Bragg, F R S, for his permission to work in his laboratory and for his helping advice concerning the research The writer is indebted to Dr J Brentano and Dr S Tolansky for experimental advice and to Dr R Peierls for his help in the theoretical discussions

SUMMARY

An optical method of measuring the concentration distribution round a growing crystal is described The flow of molecules at every point of the crystal surface can be determined The results obtained on sodium chlorate are

1—The concentration is not constant all over the crystal surface, but is highest at the corners

2—One has to assume the existence of an adsorbed layer on the crystal surface, and that a lateral flow of molecules takes place in this layer

3—There is a hitherto unexplained effect influencing the transition of molecules from the solution into the different parts of the adsorbed layer The influx is smaller at the corners than at the centres of the faces, although the concentration of the solution is higher on the corners

REFERENCES

Brummer, E 1904 *Z phys Chem* **47**, 56
Jenkins, J D 1925 *J Amer Chem Soc* **47**, 903
Le Blanc, M and Schmandt, W Z 1911 *Z phys Chem* **77**, 614
— — 1913 *Z phys Chem* **86**, 334
Marc, R 1905 *Z phys Chem* **61**, 385
— 1908*a* *Z phys Chem* **67**, 470
— 1908*b* *Z phys Chem* **68**, 104
— 1910 *Z phys Chem* **73**, 685
— 1911*a* *Z phys Chem* **75**, 710
— 1911*b* *Z phys Chem* **76**, 584
— 1912 *Z phys Chem* **79**, 71
— 1913 *Z phys Chem* **81**, 641
Miers, H A and Isaac, F 1906 *J Chem Soc (Trans)*, **89**, 413
Nernst, W 1904 *Z phys Chem* **47**, 52
Noyes, A A and Whitney, W R 1897 *Z phys Chem* **23**, 689
Roller, P S 1932 *J Phys Chem* **36**, 1202
Tolansky, S and Lee, E 1936 *J Sci Instrum* **13**, 261
Volmer, M 1932 *Trans Faraday Soc* **28**, 359

Studies of Region E of the Ionosphere

By J E Best, Ph D, F T Farmer, Ph D and
J A Ratcliffe, M A

(*Communicated by E V Appleton, F R S—Received 9 September 1937*)

1—Introduction

Many investigations of Region E of the ionosphere have been described, in the more detailed of which the critical-frequency technique of Appleton has been used to determine the penetration frequency of the region In the latest work this has generally involved continuous variation of frequency and photographic recording so that the finer details of the P'-f curve, relating equivalent path (P') to frequency (f), could be obtained It has usually happened, however, that these records have been obtained at considerable time intervals, such as half an hour or an hour, so that the finer details of the temporal variations in the behaviour of the region could not be followed

The present paper describes observations made in such a way as to exhibit the finer details of the temporal variations The Breit and Tuve pulse method was used, in conjunction with a cathode-ray oscillograph employing a time-sweep in the usual way at the receiver The oscillograph, was, however, observed visually, and the observer could alter the frequency of the transmitter (situated about 1 mile away) by means of a remote control In this way it was possible to keep a continuous detailed watch on the temporal variations of any particular feature of the P'-f curve The observations were made in Cambridge during the second half of 1936

Until very recently it has not been clear what should be taken as the critical penetration frequency for Region E, particularly when waves are reflected simultaneously from Regions E and F over a considerable range of frequency Appleton and his collaborators (1935, 1936) have recently described in detail the form of P'-f curve obtained by photographic recording under these conditions, and a sketch of an idealized P'-f curve constructed from their examples is given in fig 1, where the thickness of the lines is meant to indicate the amplitudes of the reflected waves Appleton supposes that the frequencies marked c and c' are the true penetration frequencies for Region E, and that the frequency a at which the F echo is first seen as the frequency is increased and the frequency b at which the E echo is last seen, are not, in general, simply related to the penetration frequency of the region P'-f curves observed in the summer of 1936 were usually of the type shown in fig 1, but in the winter curves of the type shown in fig 2, exhibiting clearly marked critical frequencies, were usually observed

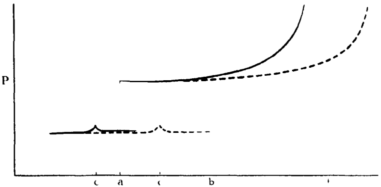


Fig 1—Idealized P' f curve for summer Continuous line represents ordinary wave, dashed line represents extraordinary wave

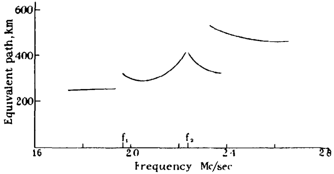


Fig 2—P' f curve for winter Ordinary wave

When using the visual method of observation, as explained above, for the purpose of repeating observations rapidly, it was found that in the summer the small changes in equivalent path, corresponding to the critical frequencies c and c', were not easily observed, and it was more natural to observe the frequencies a and b at which the amplitudes of the two echoes changed rapidly as the frequency was altered It is convenient to speak of these two frequencies as "threshold frequencies" In § 2 an account is given of the results of observations of these threshold frequencies in the summer, and particularly of their variation with time On some occasions they were found to vary very smoothly, we call these occasions ionospherically quiet and they are discussed in § 2*a* Occasions on which the threshold frequencies varied rapidly and irregularly are called disturbed, and are discussed in § 2*b* The opportunity was taken of making as many observations as possible during thunderstorms, in view of the fact that it has often been suggested that thunderstorms may influence Region *E* The summer of 1936 presented a very favourable opportunity for studying this question These observations are discussed in § 4

In the winter the visual method of observation is suited for following the changes of the true critical frequencies marked f_1 and f_2 in fig 2 Observations of these frequencies, extending over 24 hours, were carried out on a series of winter days and from the results a mean curve for a "quiet" winter day was constructed This curve is discussed in § 3*a* In that section also the true critical frequencies observed in winter are compared with the threshold frequencies observed in summer

2—SUMMER OBSERVATIONS

a—Quiet Days

Observations show that the threshold frequency *a* is usually fairly well marked, frequency *b* is sometimes clear-cut, and determinable with considerable accuracy, but on other occasions is very vague and difficult to determine When it is not clear-cut it is often found to be varying rapidly (large changes in 15 min), and the indeterminancy in the observations is not large compared with the changes occurring naturally in 5 or 10 min In determining these "threshold" frequencies there is no sudden drop of the amplitude to zero, and it might be thought that the determination is very vague Experience has shown that, by comparing the amplitude of the echo under consideration with that of one for which the amplitude is not changing rapidly with frequency, it is nearly always possible to assign a "threshold" frequency within fairly narrow limits Five separate observers obtained concordant results, and if one took over the observing from another there was no apparent break in the type of phenomenon observed

On a few special occasions the threshold frequencies were found to behave extremely smoothly, and to be very clearly marked Detailed observations were made on these occasions, and led to results of the type shown in fig 3, which represents the observations of 11 July 1936 We have not observed many occasions when the threshold frequencies varied so smoothly or were so clearly marked as those represented in this figure We suggest that under conditions such as those of fig 3, the two threshold frequencies represent the penetration of Region E by the ordinary and the extraordinary waves respectively, so that for intermediate frequencies the extraordinary wave is reflected from Region E and the ordinary from Region F * On this supposition we should not expect to observe threshold frequencies separated by smaller differences than those of fig 3 except in the case of very strong absorption, or very weak signals, when the extraordinary wave (threshold b) might not be clearly marked This has, in fact, been found to be the case, all the really well-marked and smoothly varying threshold frequencies have

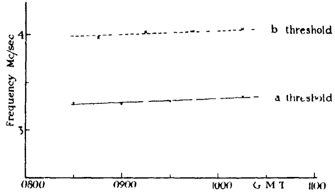


Fig 3—Observations of 11 July 1936 Very quiet conditions

behaved similarly to those of fig 3, while occasionally at midday in summer (strong absorption) the a- and b thresholds have coincided The many occasions when the separation between the a- and b-frequencies is greater than that shown in fig 3 are discussed separately in § 2b

If the two threshold frequencies a and b shown in fig 3 are coincident with the penetration frequencies for the ordinary and the extraordinary waves, then the following relation should hold between them (Appleton and Builder, 1933)

$$\frac{f_b^2}{f_a^2} = \frac{1}{1 - f_H/f_b}, \tag{1}$$

where

$$f_H = \frac{He}{2\pi mc},$$

H = earth's magnetic field, e = electronic charge, m = electronic mass, c = velocity of light Taking $H = 0{\cdot}443$ Gauss at the level of Region E (Appleton 1934) and inserting the mean observed value of f_b from fig 3 into

---

* Referring to fig 1, this means that the points a and c coincide, as also do the points b and c'

the right-hand side of (1), we calculate $f_b^2/f_a^2 = 1·45$, whereas the mean observed value of this ratio is 1·48. Appleton and Builder (1933) have pointed out that the closeness of agreement between the theory and this type of experiment means that the reflexion conditions are very accurately "quasi-transverse", which in our case requires the collisional frequency at the level of reflexion to be less than $6·5 \times 10^5$ c./sec. The agreement also shows that the charged particles in Region E responsible for the deviation of the wave are predominantly electrons rather than ions. Berkner and Wells (1934) have shown how to modify equation (1) to include the case where there are an appreciable number of ions of molecular mass in addition to electrons. Using their expression it is possible to use the nine pairs of observed frequencies plotted in fig. 3 to deduce the magnitude of the quantity

$$r = \frac{N_i/m_i}{N_e/m_e},$$

where N_i and N_e are the numbers per c.c. of ions and electrons of mass m_i and m_e. We deduce values of r lying between $+0·068$ and $-0·061$, with a mean value of $-0·005$. The negative value is meaningless, and is presumably due to experimental error, and we interpret our result as indicating that the number of ions of molecular mass is so small that it cannot be measured by this method. If we take our extreme value $r = 0·068$ and assume

$$m_i/m_e = 1840 \times 32$$

for molecular oxygen ions, we find $N_i < 4000\,N_e$.* It is to be noted that if the b-threshold frequency as observed is a little greater than the true penetration frequency for the extraordinary wave (as it often is under disturbed conditions), then the deduced value of r will be too low.

b—*Disturbed Conditions*

We now turn our attention to the occasions on which the threshold frequencies do not vary smoothly with time; we shall call these occasions "disturbed". Fig. 4 shows the results of observations made under disturbed conditions, and should be compared with fig. 3, which is characteristic of a very quiet period.† Inspection of fig. 4 reveals the following points, which are always found to characterize disturbed conditions.

(i) The difference between the threshold frequencies is often greater than the "magneto-ionic" difference observed under quiet conditions.

---

* We notice that the ions may far outnumber the electrons, although on account of their greater inertia their effect on electromagnetic waves may be negligible.

† The occasional coincidence of the a and b frequencies in fig. 4 presumably indicates that the extraordinary wave was strongly absorbed.

(ii) The b-threshold frequency is more variable than the a-threshold

(iii) The b-threshold frequency may increase rapidly by a factor of about two in a period of about ten minutes to half an hour, and usually decreases again to a more or less "normal" value in a period of the order of quarter of an hour

There is, of course, no sharp dividing line between the quiet conditions discussed in § $2a$ and the disturbed conditions exemplified in fig 4, or between these disturbed conditions and those abnormally disturbed conditions which have been styled by different workers "abnormal ionization", "sporadic ionization" or "intense E" There is such a variety in the possible behaviour of Region E that it is somewhat difficult to decide to which of these conditions the published work of other investigators refers We think, however, that most of the published work referring to "abnormal", "sporadic', or ,"intense" ionization refers to conditions which are considerably more disturbed than those with which we here deal

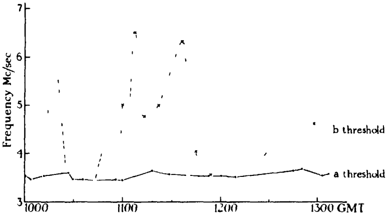


Fig 4—Observations of 3 August 1936 Disturbed conditions

In discussing observations of the type illustrated in fig 4 we first consider the observation (i) in the above list, that the separation between the a- and b-threshold frequencies is often greater than the magneto-ionic value Several workers* have suggested that this phenomenon represents a case of partial reflexion from a region with a sharp boundary The a-threshold is then supposed to correspond to the "total reflexion' of the waves and is determined by the maximum value of the ionization density, and the b-threshold is supposed to be determined by the sharpness of the lower boundary of the region We see two objections to this hypothesis In the first place, in order to explain the fact that the b-frequency varies rapidly over a much wider range than the a-frequency we must assume that there may be great changes in the sharpness of the lower boundary of the layer, while changes in the maximum density of the layer are comparatively small Now we believe that on quiet days Region E approximates very closely to a simple Chapman region, with H of the order of 10 km , since the absorption can be very completely explained on this hypothesis * It can be shown that a Chapman region of this kind has nowhere an ionization gradient sufficiently sharp to produce appreciable partial reflexion Let us assume that Region E is like a Chapman region on quiet days, and let us consider how it might be modified to give appreciable partial reflexion on

---

* E g Kirby and Judson (1935), Gilliland (1935)

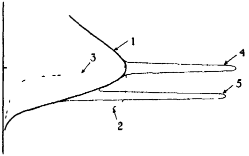


FIG 5—Chapman region with possible modifications to account for observed facts

disturbed days In fig 5, curve 1 represents the ionization distribution in a Chapman region, and curves 2, 3, 4 and 5 represent possible modification of this distribution to give sharper gradients of ionization Curves 4 and 5 would not fit in with the observations recorded in fig 4, since they would involve a considerable change in the a-threshold frequency which depends on the maximum density It is probable that ionization distributions of this kind are responsible for the extreme cases of disturbance known as intense (abnormal, or sporadic) ionization Curves 2 and 3 would fit in with the observations, since the maximum ionization density is unchanged It can, however, be shown that the absorption of a wave traversing a Chapman region occurs chiefly at the point of inflexion of the curve, and it is just at this point that the curve has been modified, a rough calculation shows that if the gradients in curves 2 and 3 are made great enough to give appreciable partial reflexion then the absorption would be very noticeably altered Measurements of the Region E absorption of Region F echoes during

---

* See a paper in the press by Best and Ratcliffe

"disturbed" periods in Region *E* have shown no variations of absorption of the magnitude expected We conclude that it is impossible to devise a "sharp-boundary" distribution of ionization which will explain simultaneously the observed changes in the *a*- and *b*-threshold frequencies and the absorption of echoes traversing Region *E*

We would also criticize the partial reflexion hypothesis on the ground that it requires the lower edge of the region to be sharp *compared with distances of the order of one wave-length* (100 m) This appears to us to be most unlikely, because an ionized region with the required gradient of ionization would rapidly have its gradient reduced by the process of diffusion This process is discussed more fully later

For the reasons given above we do not consider that the Sharp boundary theory can give a reasonable explanation of disturbed conditions, we think a more satisfactory explanation can be given along the following lines We suppose that Region *E*, under disturbed conditions, is not uniformly stratified in horizontal planes, but contains irregularities of the nature of electronic "clouds" These "clouds" are regions of more intense ionization immersed in a region which, on the average, is like a Chapman region This suggestion has been made before,* and it has been shown that it would explain the facts of lateral deviation and of the fading of single echo pulses It has been pointed out that, on this picture, an ionospherically reflected wave does not travel along a single ray but part of the energy is returned from a series of scattering centres distributed over an area which may have linear dimensions (in the horizontal direction) of the order of 30 km The scattering centres have dimensions of the order of one wave-length (100 m) This hypothesis would explain the observations on the "disturbed" Region *E* very simply The *a*-threshold frequency would represent ordinary waves which could just penetrate the most feebly ionized portion of the region, while the *b*-threshold would represent the penetration of the most densely ionized portion † It is at once evident that there would be no definite relation between the two threshold frequencies if some clouds of especially dense ionization were suddenly produced the *b*-frequency would increase, without change in the *a*-frequency, if a large number of such clouds were produced they would eventually, by diffusion, add to the general ionization of the region, and would cause a (smaller) increase in the *a*-threshold frequency This would explain a curve of the form shown in fig 4 On our picture when *F* and *E* echoes occur together the *F* echoes may be thought of as having gone "through the holes" in Region E, while the E echoes are reflected "from the opaque parts" *

---

* E g Eckersley (1932), Watson Watt (1933), Ratcliffe and Pawsey (1933)

† In the absence of absorption this frequency would represent penetration by an extraordinary wave, but in presence of strong absorption by the ordinary wave

We have pointed out that the variations in the b-threshold frequency occur fairly rapidly, in times of the order of 15 min This might be due to the fact that electronic "clouds" were produced and dissipated in times of this order of magnitude, or that a "cloudy region" drifted past the observing point in a time of this order If we take the value deduced later (§ 3a) for the recombination coefficient in the day-time we find that a cloud having ionization density twice that of Region E at midday would have its density reduced to that of Region E in about 10 min if all ionizing influences were suddenly removed

Small clouds of electrons would not disappear simply as a result of recombination, but would smooth themselves out by a process of diffusion To find the order of the time required we consider a plane boundary having an ionization gradient such that the electron density drops to e^{-1} of its original value in a distance of 100 m We assume that, at a level of 120 km (average level of Region E) the mean free path of the ions† is 10 cm and the ionic velocity is 5×10^4 cm /sec ,‡ and then ordinary diffusion theory shows that the electron gradient would decrease with a time constant of about 10 min This calculation is also of interest in setting a limit to the possible sharpness of the lower edge of Region E It shows that a region sufficiently sharp to reflect by the process of partial reflexion would very soon lose its steep gradient by a diffusion process

It is difficult to envisage a mechanism which might be responsible for the production of clouds of electrons They might be due to small groups of ionizing particles incident from outside the atmosphere If this is the case it is pertinent to ask why these particles produce their ionization at the level of the already existing (Chapman) layer, the answer may be that in the existing layer there are already a large number of negative molecular ions or of excited atoms or molecules from which the electrons are easily removed by the incident particles

Experiments of the same type as those described above have shown that, although there are small rapid and irregular changes in Region F ionization, their magnitude does not at all approach those observed in Region E, and in particular the small changes which are observed in Region F appear

---

* An analogy in sound would be the double echo produced by a line of trees standing some distance in front of a vertical cliff face

† The electrons cannot diffuse more rapidly than the ions if they did the resulting space charge would prevent further diffusion

‡ These values accord with the suggestions put forward by Martyn and Pulley (1936), which we have found to be in agreement with absorption measurements

to be unrelated to those observed in Region E In connexion with this observation it is important to note that on account of the larger mean free path and possibly greater temperature the rate of diffusion of an electron cloud in Region F would be considerably more rapid than that in Region E, so that the clouds might diffuse as rapidly as they were formed

3—Diurnal variations

a—Winter Observations

We have previously pointed out that in winter it was often possible to follow the variations of the true critical frequencies marked f_1 and f_2 in fig 2, though sometimes these were not clear and only "threshold" frequencies could be observed A series of experiments was made on ionospherically quiet days near 16 December 1936 with the purpose of investigating in detail the diurnal variation of the critical frequencies

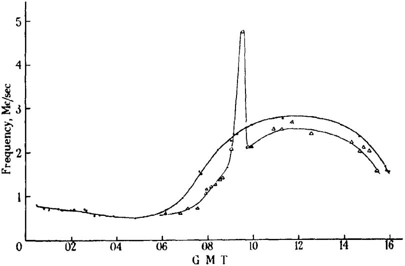


Fig 6—Observations of 18 December 1936 ● Frequency f_2, △ Frequency f_1 (see fig 2)

On these days observations were made, as far as possible, during the whole 24 hours, but, owing to the overlap of the critical frequencies with the broadcast band, it was not possible to observe during the period 1530–2400 G M T Fig 6 shows the results obtained on 18 December 1936, and is typical of the behaviour on these days The triangles and dots plotted in this figure refer to the critical frequencies marked f_1 and f_2 in fig 2 It was not possible to observe these frequencies separately during the night hours Between 0900 and 1000 G M T there are shown some abnormal readings representing a change to the "threshold frequency" type of behaviour, after one hour the readings became more normal Whenever short-lived abnormalities of this kind were observed they were neglected in drawing smooth curves through the points

From a set of observations such as those shown in fig 6 a smooth curve was drawn for each separate day Fig 7 shows the kind of agreement between the different days In this figure the individual points are taken from smoothed curves plotted for each day as in fig 6 The good agreement between the different days lends support to the assumption that these days can be used in investigating the behaviour of the ionosphere on quiet days Smooth curves have been drawn through the points in fig 7, and we shall take these curves to represent the average quiet day behaviour near 16 December

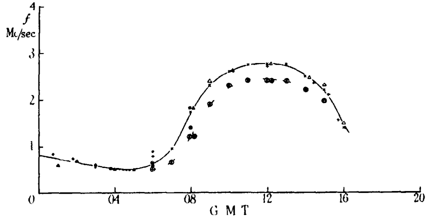


Fig 7—Points from smoothed curves similar to those of fig 6
● ⊙ 15 December, + ⊕ 16 December, × ⊗ 17 December, △ ◬ 18 December

In considering the results theoretically we first take the night-time portion of the curve and consider the rate of electron loss Appleton and Naismith (1935) have shown that at midday the loss of electrons is due to recombination, and from the night-time curves the coefficient of recombination α can be calculated We first notice that the ionization begins to increase in the early morning about 3 hr before sunrise This would not be expected on a simple theory, even when the curvature of the earth is taken into account, and it is difficult to see how it could be due to diffusion of electrons laterally from the illuminated part of the atmosphere In order to calculate a value for α we consider the interval from sunset to the time of the minimum before sunrise and find that the experimental curve can be fitted reasonably well with a value of $\alpha = 4 \times 10^{-9}$ This is somewhat greater than the theoretical value for recombination of electrons and positive ions at those pressures for which α is independent of pressure Chapman (1931*b*) calculates α to be of the order 10^{-10} for these conditions It is of interest to note that if electron attachment is assumed so that $\frac{dN}{dt} = -\beta\nu N$ (β = probability of attachment, ν = collisional frequency) then, taking $\nu = 10^5$ per sec we find $\beta = 5 \times 10^{-10}$ This value of β is very different from the generally accepted value of about 10^{-5} (Bailey 1925)

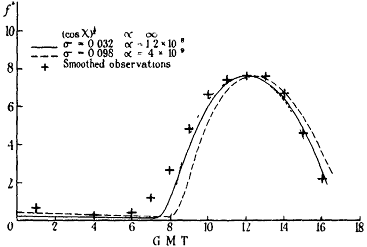


FIG 8—Comparison of theoretical curves with smoothed observations of critical frequency in winter

Turning now to the day-time portions of the curves, we first note that the winter P'-f curve shows two clean-cut jumps, one from a level of about 130 km to a level of about 150 km , and the other from 150 km to Region F (see fig 2) We first restrict ourselves to the higher critical frequency (Region E_2) represented by the upper curve in fig 7 Chapman (1931*a*) has shown that in a simple ionospheric region in which the ionization density (N) is always in equilibrium with the ionizing agency, N should be proportional to $(\cos\chi)^{\frac{1}{2}}$, where χ is the sun's zenith distance, if recombination is the cause of electron loss In fig 8 the dotted line represents the diurnal variation of $(\cos\chi)^{\frac{1}{2}}$ on 16 December, and the crosses represent the smoothed experimental values from fig 7 In fig 8 the values of f^2 are shown, since f^2 is proportional to N, and the ordinates have been adjusted so that the value of $(\cos\chi)^{\frac{1}{2}}$ at midday coincides with the observed value of f^2 Fig 8 illustrates the point, mentioned by other workers (Kirby and Judson 1935, Appleton 1935), that the observed values agree fairly closely with the curve of $(\cos\chi)^{\frac{1}{2}}$ during the period from sunrise to sunset If the observed points coincided accurately with this curve it would mean that there was equilibrium between the ionization density and the ionizing agency throughout the day It is of interest to find out how far the finite value of α would cause the ionization to depart from the equilibrium value, and hence cause the ionization curve to depart from the $(\cos\chi)^{\frac{1}{2}}$ curve This involves the solution of the equation

$$\frac{d}{dt}\{N(z,t)\} = q(z,t) - \alpha\{N(z,t)\}^2, \tag{2}$$

which relates the value of N at the height z and time t to the number $q(z,t)$ of ions produced per second at that height and time Chapman has given data from which $q(z,t)$ can be calculated for a spherical earth Equation (2) has been solved for us, for various values of α, by Mr M V Wilkes * When numerical solutions have been obtained for a series of given heights using a constant value of α it is possible to construct curves showing maximum ionization density as a function of time of day (the level at which the maximum occurs varies, of course, throughout the day) The curves to which we shall now refer in fig 8, and also those in fig 11, have been calculated in this way, and we would like to take this opportunity of expressing our thanks to Mr Wilkes for the help he has given by performing these calculations for us †

The dashed curve of fig 8 was calculated using the value of $\alpha(4\times10^{-9})$ deduced from the night-time observations It is at once obvious that the finite value of α does not cause any very large departure from the $(\cos\chi)^{\frac{1}{2}}$ curve, but, in view of the fact that the observational points represent the mean of 4 days chosen to be specially quiet, we consider that they are sufficiently accurate to indicate better agreement with the equilibrium curve than with the curve for $\alpha = 4\times10^{-9}$ We therefore try the effect of a larger α, and the continuous curve, calculated for $\alpha = 1\ 2\times10^{-8}$, shows the

---

* The calculation was performed by means of a "differential analyser" of the Bush type, kindly lent for the purpose by Professor Lennard Jones Mr Wilkes hopes to give an account of the method of calculation in a forthcoming paper

† It is to be noted that the dotted curve of fig 8, which represents the function $(\cos\chi)^{\frac{1}{2}}$, would only be expected to be proportional to N for a plane earth, whereas the continuous curve and the dashed curve have been calculated by using Chapman's data computed for a spherical earth

result This curve agrees better with the experimental points, although an even larger value of α might fit better still If we confine our attention to the interval comprised between 1 hr after sunrise and 1 hr before sunset* it appears that the results agree, to the order of experimental accuracy, with a value of α a little greater than $1 \cdot 2 \times 10^{-8}$ We have previously noticed that the decay of ionization during the night, in the interval between sunset and the early morning minimum, can be explained if we take $\alpha = 4 \times 10^{-9}$, the assumption of $\alpha = 1 \cdot 2 \times 10^{-8}$ would require the observed decay to take place in a time which is only one third of the observed time We are therefore driven to conclude that the day-time value of α is greater than the night-time value, and considerably greater than the usually accepted value for electrons recombining with positive ions at low pressures (10^{-10})

The effect of any lagging of ionization behind the variations of the ionizing agency would be expected to be even more marked in the case of eclipses, where the rate of change of ionization is greater than in the ordinary diurnal variations Examination of the Region E curves obtained by Kirby, Gilliland and Judson (1936) on the occasions of the eclipses of 31 August 1932 and 5 February 1935 shows no apparent lag, which again points to a very large value of α in this region

These conclusions about α are in accord with those which we have drawn elsewhere from a study of the absorption of waves and also agree with the views recently expressed by Appleton (1937)

Several workers have calculated the value of q (equation (2)) at the level of maximum ion-production (q_m), by writing for the midday (equilibrium) conditions $q_m = \alpha N_m^2$, α being taken as equal to about 10^{-10}, and have then proceeded to discuss the intensity of the ionizing radiation outside the atmosphere and to relate it to the radiation from the sun, considering it as a black body These calculations need revision in view of the large value of α which is here suggested

b—Summer Observations

We have already pointed out that the threshold frequencies observed in the summer are not necessarily the true critical frequencies for Region E In an attempt to relate the summer threshold frequency to the winter critical frequency, observations were made of the *a*-threshold frequency throughout an extended period on several ionospherically quiet days in summer, and in winter similar observations were made on the greater critical frequency (f_2 in fig 2) All these observations are collected in fig 9 in which the observed value of f^4 is plotted against cos χ It is seen that the summer values for a-threshold frequencies agree well with the winter values for critical frequencies and that all show a proportionality between f^4 and cos χ, this leads us to suppose that, *on the quiet days chosen*, the a-threshold frequency represents the Region E critical frequency in summer In this connexion it is to be noted that the greater critical frequency (f_2) in winter is compared with the a-threshold frequency in summer, if the smaller critical frequency had been taken the points for the 4 days near 16 December would have given the line shown dotted in fig 9

---

* Ground sunrise and sunset correspond to the intersections of the dotted curve with the axis

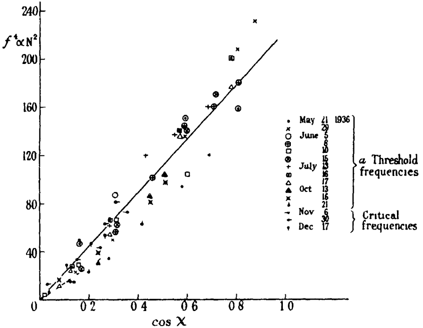


FIG 9—Observed values of critical frequency (during winter) and a threshold frequency (during summer) compared with theoretical relation $f^4 \propto \cos\chi$ The variations of χ in this figure represent partly variation throughout a given day and partly variation throughout the year

The agreement between the summer a-threshold frequency and the winter critical frequency prompts us to make a detailed investigation of the diurnal variation of the summer threshold frequency Experiments were made on several days and nights near 31 July, occasions which were ionospherically quite alone being investigated Fig 10 shows the results This figure is analogous to fig 7 for winter, the points are taken from smoothed curves for each separate day, and the smooth curve is drawn to represent an average for days near 31 July Fig 11 [analogous to fig 8 for winter] shows how the smoothed experimental results (crosses) compared with the calculated values for $(\cos\chi)^{\frac{1}{2}}$ [dotted], for $\alpha = 4 \times 10^{-9}$ (dashed) and for $\alpha = 1\ 2 \times 10^{-8}$ (continuous) It is first noticeable that the difference between the theoretical

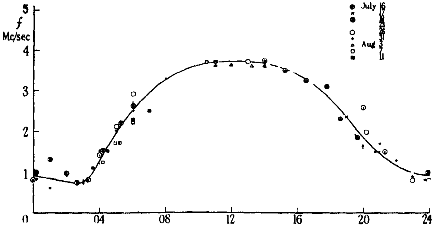


FIG 10—Diurnal variation of *a* threshold frequency near 31 July 1936

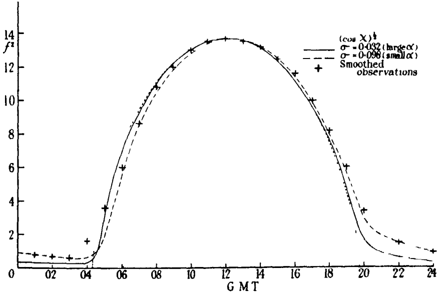


FIG 11—Comparison between theoretical curves and smoothed observations of *a* threshold frequency near 31 July 1936

curves is not so great as in winter, so that the experimental results cannot be used to discriminate between different values of α, particularly since the nature of the observations renders them less accurate than the winter ones The accuracy is, however, sufficient to show that the fit with any of the theoretical curves is fairly close An examination of the night decay leads to $\alpha = 4 \times 10^{-9}$ as in winter It is again noticeable that the ionization begins to increase some considerable time before ground sunrise, though the interval appears to be a little less than in winter

The close agreement between the observed and calculated diurnal variation of the threshold frequency, the agreement between the summer a-threshold frequency and the winter critical frequency exhibited in fig 9, and the agreement of the two threshold frequencies with the magneto-ionic theory on very quiet days, as pointed out in § 2a, all support the suggestion that *on quiet days* the a-threshold frequency represents the ordinary wave penetration frequency for Region E

4—REGION E AND METEOROLOGICAL PHENOMENA

a—*Thunderstorms*

In 1932–3 one of us in collaboration with E L C White (Ratcliffe and White 1934) made an investigation of the possible effect of thunderclouds on Region E ionization The data were obtained from records of equivalent height on a fixed frequency (P'-t records), and a "disturbed" day was identified by the occurrence of Region E reflexions at times when they would not normally occur A statistical study of the results appeared to indicate a correlation between the occurrence of thunderstorms and of abnormal Region E reflexions

Several other workers* have investigated the matter, but there does not yet seem to be any agreement between the different results Since the summer of 1936 was particularly rich in thunderstorms in south-east England the opportunity was taken for adding to the rapidly accumulating data on the subject

Observations were made on as many occasions as possible during thundery weather, including 5 days when a storm was actually in progress at Cambridge Measurements were made of the b-threshold frequency described above The observations described in previous sections provided comparison data for days of a non-thundery character For the purpose of a statistical analysis two different numbers were used to represent the amount of disturbance in Region *E* In the first case a number 0 to 3 was attributed to the day according to the degree of variability of the *b*-threshold frequency thus in the case of fig 3 the number would be 0, and of fig 4 it would be 3 In the second case the number used was the greatest *b*-threshold frequency occurring during any one period of observation In order to indicate the degree of "thunderiness" a figure on the scale 0 to 3 was derived from the daily reports of the meteorological office This figure was assigned on the basis of the location and time of occurrence of thunderstorms reported in the British Isles The higher numbers were assigned to storms occurring in south-east England during the 8 hr interval which contained the wireless observations The results are collected in Table I, which represents observations on 21 days between 1 June and 3 August and 2 days in September on which local thunder occurred

---

* A summary of work on this subject is given by Mimno (1936) and some more recent work is described by Bhar and Syam (1937) and by Appleton, Naismith and Ingram (1936)

TABLE I—1936

Thunder character figure	0	1	2	3
Mean value of the highest threshold frequency observed	5 5	5 1	6 0	5 0
Mean value of the variability figure for the threshold frequency	1 5	1 2	1 2	1 7
Number of days in the group	7	5	5	6

Since these results do not show any correlation of the type noted in the 1932 experiments, the recent results have also been analysed in a different way in order to provide a basis for direct comparison with the fixed frequency data of the earlier experiments For this analysis the days of observation were divided according to whether the highest threshold frequency was greater or less than the fixed value 5 Mc /sec The following figures were obtained (Table II)

TABLE II—1936

	Mean thunder figure
Days with threshold frequency greater than 5 Mc /sec (10 days)	1 2
Days with threshold frequency less than 5 Mc /sec (13 days)	1 6

All three methods of analysis appear to indicate no correlation between the disturbed Region *E* conditions and the presence of thunder Since the old (1933) observations gave the opposite result it is important to reexamine them, in a way as nearly as possible the same as that used for the 1936 results

For this re-examination some later results obtained in the second half of 1933 and 1934 by the P'-t recording method of 1932–3 were included in the data The thunder index figure was assigned as explained for 1936, and Region E was considered to be "disturbed" during the night if the b-threshold frequency rose above 4 Mc /sec Table III shows the result of this

TABLE III—1932–4

	Mean thunder figure
Nights with threshold frequency greater than 4 Mc /sec (77 nights)	0 55
Nights with threshold frequency less than 4 Mc /sec (83 nights)	0 45

analysis The difference between the thunder figures is not great enough to be considered significant, and we conclude that when the results are analysed in this way they do not show any correlation The difference in the results of this method of analysis and the older method is due partly to a stricter selection of thunderstorms so that more weight is given to those occurring in south-east England near the time of the observations, and partly to the inclusion of the newer data for 1934 and the second half of 1933 Taking our results as a whole there does not now seem to be much evidence that thunder-clouds can influence Region E ionization

b—Barometric Pressure

Martyn and Pulley (1936) have drawn attention to an interesting correlation between the occurrence of nocturnal Region E ionization and the barometric pressure observed at the ground next morning Their observations refer to Australia, and they point out that a similar correlation has not yet been reported from the northern hemisphere It is of interest to examine the observations made during the 1932–4 series of experiments in order to see whether there is any correlation The conditions laid down for the occurrence of nocturnal ionization were that the b-threshold frequency should rise above 4 Mc /sec or that the a-frequency should rise above 2 Mc /sec for a period of more than 15 min in each case These conditions are approximately the same as those used by Martyn and Pulley The barometric pressure as observed in Cambridge at 9 a m on the morning following each night's observations was noted, and the mean pressures for the nights showing nocturnal Region *E* ionization and for those showing no such ionization are given in the following table

	Mean value of barometric pressure	Standard deviation of the mean
Nocturnal *E* ionization present (131 nights)	1017 8 m b	0 55
Nocturnal *E* ionization absent (160 nights)	1016 0 m b	0 31

The difference of the mean is 1 8 and comparison with the standard deviations indicates that the difference has no significance It appears therefore that the conditions are not the same in England as they are in Australia

In conclusion we wish to express our thanks to Professor Appleton for several helpful discussions and particularly for advice in connexion with the presentation of the paper, to the Department of Scientific and Industrial Research for grants which enabled us to carry out this work, and to Messrs K G Budden and K Weekes for valuable help with the observations Our great indebtedness to Mr M V Wilkes, who made the calculations of § 3, has already been mentioned

SUMMARY

Investigations of Region *E* of the ionosphere have been made using a Breit and Tuve pulse sender of variable frequency By using visual observation at the receiver and a remote control of the sender it was possible to follow the variations of frequencies characteristic of the ionosphere even when ionospheric conditions were changing rapidly

In the winter it was possible to observe true critical frequencies for Region *E* and a curve is given (fig 7) which represents the mean of observations taken on "ionospherically quiet" occasions near 16 December 1936 From this curve it is deduced that the night-time value of the recombination coefficient is 4×10^{-9}, whereas the day-time value is of the order of three times greater

In the summer it was more natural to observe certain "threshold frequencies" at which the amplitudes of the Region *E* and Region *F* echoes underwent marked changes Reasons are given for supposing that, on ionospherically quiet days, one of these threshold frequencies approximates closely to the true critical frequency of Region *E* The diurnal variation of

this threshold frequency was studied for "quiet" days Observations of the threshold frequencies on days which were ionospherically disturbed can be explained on the supposition that Region *E* is then not uniformly stratified in horizontal planes but contains "clouds" with electron density greater than that of their surroundings The simultaneous reflexion of waves from Regions *E* and *F* over a considerable range of frequencies is explicable on this hypothesis The conditions here considered are supposed to be intermediate between those which we call "quiet" and those which are characterized by the occurrence of "abnormal", "intense", or "sporadic" Region *E* ionization

An investigation of Region *E* during thunderstorms and thundery weather in 1936 showed no correlation between the storms and the behaviour of the region

An examination of some data obtained between 1932 and 1934 shows no correlation, such as Martyn and Pulley (1936) found in Australia, between the occurrence of nocturnal Region *E* ionization and barometric pressure

REFERENCES

Appleton 1933 *Proc Phys Soc* **45**, 673
— 1934 *Nature, Lond* **133**, 793
— 1935 *Phys Soc Reports on Progress in Physics* p 129
1937 *Proc Roy Soc* **162**, 451
Appleton and Builder 1933 *Proc Phys Soc* **45** 208
Appleton and Naismith 1932 *Proc Roy Soc* **137**, 36
— — 1935 *Proc Roy Soc* **150**, 685
Appleton, Naismith and Ingram 1936 *Philos Trans* A, **236**, 191
Bailey 1925 *Phil Mag* **50**, 825
Berkner and Wells 1934 *Proc Inst Radio Engrs, N Y* , **22**, 1120
Bhar and Syam 1937 *Phil Mag* **23**, 513
Chapman 1931*a* *Proc Phys Soc* **43**, 26 and 483
— 1931*b* *Proc Roy Soc* **132**, 363
Eckersley 1932 *Proc Inst Electr Engrs* **71**, 405
Gilliland 1935 *Proc Inst Radio Engrs, N Y* **23**, 1076
Kirby, Gilliland and Judson 1936 *Proc Inst Radio Engrs, N Y* , **24**, 1027
Kirby and Judson 1935 *Proc Inst Radio Engrs, N Y* , **23**, 735
Martyn and Pulley 1936 *Proc Roy Soc* **154**, 455
Mimno 1937 *Rev Mod Phys* **9**, 1
Ratcliffe and Pawsey 1933 *Proc Camb Phil Soc* **29**, 316
Ratcliffe and White 1934 *Proc Phys Soc* **46**, 107
Watson Watt 1933 *Proc Roy Soc* **141**, 715

Stability of Polyatomic Molecules in Degenerate Electronic States

II—Spin Degeneracy

By H A Jahn

Davy-Faraday Laboratory, The Royal Institution

(*Communicated by Sir William Bragg, O M , P R S —Received 22 September* 1937)

Introduction

In a previous paper (Jahn and Teller 1937) the following theorem was established A configuration of a polyatomic molecule for an electronic state having orbital degeneracy cannot be stable with respect to all displacements of the nuclei unless in the original configuration the nuclei all lie on a straight line The proof given of this theorem took no account of the electronic spin, and in the present paper the justification of this is investigated An extension of the theorem to cover additional degeneracy arising from the spin is established, which shows that if the total electronic state of orbital and spin motion is degenerate, then a non-linear configuration of the molecule will be unstable unless the degeneracy is the special twofold one (discussed by Kramers 1930) which can occur only when the molecule contains an odd number of electrons The additional instability caused by the spin degeneracy alone, however, is shown to be very small and its effect for all practical purposes negligible The possibility of spin forces stabilizing a non-linear configuration which is unstable owing to orbital degeneracy is also investigated, and it is shown that this is not possible except perhaps for molecules containing heavy atoms for which the spin forces are large Thus whilst a symmetrical nuclear configuration in a degenerate orbital state might under exceptional circumstances be rendered stable by spin forces, it is not possible for the spin-orbit interaction to cause instability of an orbitally stable state

1—General theorem for molecules with spin

Just as before we must see how the symmetry of the molecular framework determines whether the energy of a degenerate electronic state with spin depends linearly upon nuclear displacements This is again determined by the existence of non-vanishing perturbation matrix elements which are linear in the nuclear displacements These matrix elements are integrals involving the electronic wave functions with spin and the nuclear displacements, and we deduce as before from their transformation properties whether for a given molecular symmetry they can be different from zero

We will see that if the molecule contains an even number of electrons then the transformation properties of the integrals involving the electronic wave functions with spin are identical with those of the corresponding case involving only orbital wave functions Thus the previous theorem proven for orbital degeneracy holds also for spin degeneracy if there is an even number of electrons For an odd number of electrons with spin, however, the transformation properties of the integrals differ owing to the two-valuedness of the spin wave functions A special investigation is needed here and shows that in the case of an odd number of electrons with spin electronic states with twofold degeneracy can exist which are stable with respect to all nuclear displacements This is in accordance with the results of Kramers and Wigner, who have shown that such twofold degeneracy cannot be split by any electrical forces For electronic states which are more than twofold degenerate, however, we show that there always exist non-vanishing matrix elements which are linear in at least one set of non-totally symmetrical nuclear displacements, unless all the nuclei in the molecule lie on a straight line Thus even if the degeneracy arises from the spin a non-linear polyatomic molecule cannot be stable in a degenerate electronic state, excepting this degeneracy be the special twofold one of Kramers and Wigner

2—MATHEMATICAL FORMULATION AND GROUP-THEORETICAL CONSIDERATIONS

The linear matrix elements whose transformation properties we have to investigate are integrals of the form

$$\int \phi_\rho^* V_r \phi_\sigma d\tau,$$

where the ϕ_ρ, ϕ_σ are electronic wave functions with spin and the V_r are functions of the electronic space co-ordinates alone (A star is used in the following throughout to denote the conjugate complex) The indices ρ, σ refer to independent wave functions of the degenerate energy level and we denote as before the representation of the molecular symmetry group subtended by these functions by Φ The index r refers to the independent normal nuclear displacements and we denote the representation subtended by the V_r by V In the case of orbital degeneracy the wave functions ϕ_ρ could be chosen real and the integrals transformed according to the representation $V[\Phi^2]$, where $[\Phi^2]$ denotes the representation of the symmetrical product of Φ with itself It is our purpose to show that in the case of an even number of electrons with spin the integrals still transform according to the representation $V[\Phi^2]$, whilst in the case of an odd number of electrons with spin they transform according to the representation $V\{\Phi^2\}$, where $\{\Phi^2\}$ denotes the representation of the antisymmetrical product of Φ with itself

To establish this we make use of the properties of the symmetry operation of time reversal which have been investigated by Wigner (1932) He showed that for electrons with spin this operation has different properties according as the number of electrons in the system is even or odd, and it is this fact which gives rise to the different behaviour of the integrals in the two cases We will further make use of the results of the fundamental work of Frobenius and Schur (1906) on the real representations of finite groups They investigate the properties of representations which leave a certain form G invariant, and we will show that this form G has the same properties as the matrix (K) representing the operation of time reversal This enables us to write down almost at once the relations we need from the results of Frobenius and Schur

Wigner shows that the operation K of reversing the time is a non-linear operator for the wave functions, acting as follows upon a linear combination of any two wave functions ϕ and ψ

$$K(a\phi+b\psi) = a^*K\phi+b^*K\psi$$

He further shows that the operator K commutes with the operator R of any spatial rotation or reflexion (i e symmetry operation)

$$KR = RK$$

Using any complete set of independent wave functions of an energy level the operator K can be represented by a matrix which we will denote by (K) Owing to the non-linearity of K the operator K and its matrix (K) do not obey the same commuting rules Thus if

$$R\phi_\rho = \sum_\sigma R_{\sigma\rho}\phi_\sigma$$

and

$$K\phi_\rho = \sum_\sigma (K)_{\sigma\rho}\phi_\sigma,$$

we have

$$KR\phi_\rho = \sum_\sigma R^*_{\sigma\rho}K\phi_\sigma = \sum_\sigma\sum_\tau R^*_{\sigma\rho}(K)_{\tau\sigma}\phi_\tau$$

$$RK\phi_\rho = R\sum(K)_{\sigma\rho}\phi_\sigma = \sum\sum(K)_{\sigma\rho}R_{\tau\sigma}\phi_\tau,$$

from which we deduce, since $KR = RK$

$$\sum_{\sigma} R^*_{\sigma\rho}(K)_{\tau\sigma} = \sum_{\sigma}(K)_{\sigma\rho} R_{\tau\sigma}$$

or $$\{(K)\, R^*\}_{\tau\rho} = \{R(K)\}_{\tau\rho},$$

i e $$(K)\, R^* = R(K)$$

(There is no need to distinguish between the operator R and its corresponding matrix since R is a linear operator)

Now Wigner shows that if the number of electrons is even then

(a) $$K^2 = +1,$$

whilst if the number of electrons is odd

(b) $$K^2 = -1$$

The relation $$(K)R^* = R(K)$$

gives $$(K)^2\, R^*(K) = (K)R(K)^2,$$

from which we deduce in both cases

$$R^*(K) = (K)\, R$$

or (c) $$R'(K)\, R = (K),$$

where R' denotes the transposed matrix and we have $R^* R' = E$, since R is unitary

The relations (a), (b) and (c) are identical with those postulated by Frobenius and Schur for their invariant form G, and we can hence make use of their results directly by identifying our (K) with their G They show that in case (a) i e when $(K)^2 = E$ the independent vectors subtending the representation (or the wave functions of the energy level) can be so chosen that (K) is identical with the unit matrix E and the matrices R (and consequently the wave functions) can at the same time be chosen all real Thus in the case of an even number of electrons with spin the results are the same as when there is no spin present and the linear matrix elements

$$\int \phi^*_\rho \phi_\sigma V_r\, d_\tau = \int \phi_\rho \phi_\sigma V_r\, d_\tau = \tfrac{1}{2}\int (\phi_\rho \phi_\sigma + \phi_\sigma \phi_\rho) V_r\, d\tau$$

subtend the representation $V[\Phi^2]$ as described in the previous paper In case (b), where $(K)^2 = -E$, Frobenius and Schur show that the independent vectors subtending the representation Φ, which they show must now have an even number of dimensions (say $2n$), can be chosen so that

First, (K) is the matrix

$$\begin{array}{|c|c|}\hline O & -E \\ \hline E & O \\ \hline\end{array},$$

where E and O denote the unit and zero n-dimensional matrices respectively We may express this by putting

$$K\phi_\rho = \tilde{\rho}\phi_{-\rho},$$

where $\tilde{\rho} = +1$ if ρ is positive and $= -1$ if ρ is negative (Positive values of ρ number the first n rows and columns of the representation, negative values the last n rows and columns)

And secondly, the matrices $R_{\sigma\rho}$ for the rotations or reflexions R of the group

$$\phi'_\rho = \sum_\sigma R_{\sigma\rho}\phi_\sigma$$

can be so chosen that

$$R = \begin{array}{|c|c|}\hline A & B \\ \hline -B^* & A^* \\ \hline\end{array},$$

or

$$R^*_{\sigma\rho} = \tilde{\sigma}\tilde{\rho}R_{-\sigma\ -\rho}$$

Now from Wigner's results we deduce that since the V_r in his terminology are real (they do not involve the spins) we have the relation

$$(\phi_\rho, V_r\phi_\sigma) = (K\phi_\rho, KV_r\phi_\sigma)^*,$$

i e

$$\int\phi^*_\rho V_r\phi_\sigma\,d\tau = \int(K\phi_\sigma)^* V_r(K\phi_\rho)\,d\tau$$

But from the above relations we find

$$\int(K\phi_\sigma)^* V_r(K\phi_\rho)\,d\tau = \int\tilde{\sigma}\phi^*_{-\sigma}V_r\tilde{\rho}\,\phi_{-\rho}\,d\tau$$

Thus the integrals satisfy the relation

$$\int\phi^*_\rho\,\phi_\sigma V_r\,d\tau = \tilde{\rho}\tilde{\sigma}\int\phi^*_{-\sigma}\phi_{-\rho}V_r\,d\tau$$

This relation can be written in the form

$$\int \phi^*_{-\rho}\,\phi_\sigma V_r\,d\tau = -\tilde{\rho}\tilde{\sigma}\int \phi^*_{-\sigma}\,\phi_\rho V_r\,d\tau$$

or

$$-\tilde{\rho}\int \phi^*_{-\rho}\,\phi_\sigma V_r\,d\tau = \tilde{\sigma}\int \phi^*_{-\sigma}\,\phi_\rho V_r\,d\tau$$

Introducing the abbreviation

$$\psi_\rho = -\tilde{\rho}\,\phi^*_{-\rho}$$

we then have

$$\int \psi_\rho\,\phi_\sigma V_r\,d\tau = -\int \psi_\sigma\,\phi_\rho V_r\,d\tau = \tfrac{1}{2}\int (\psi_\rho\,\phi_\sigma - \psi_\sigma\,\phi_\rho)\,V_r\,d\tau$$

Hence if we can show that the ψ_ρ transform according to the same representation Φ as the ϕ_σ we can conclude that the integrals transform according to the representation $V\{\Phi^2\}$

Now from the transformation formulae

$$\phi'_\rho = \sum_\sigma R_{\sigma\rho}\,\phi_\sigma$$

follows, since

$$R^*_{\sigma\rho} = \tilde{\sigma}\tilde{\rho}R_{-\sigma\ -\rho},$$

$$\begin{aligned}(\phi^*_\rho)' &= \sum_\sigma R^*_{\sigma\rho}\,\phi^*_\sigma \\ &= \tilde{\rho}\sum_\sigma \tilde{\sigma}R_{-\sigma\ -\rho}\,\phi^*_\sigma \\ &= \tilde{\rho}\sum_\sigma (-\tilde{\sigma})\,R_{\sigma\ -\rho}\,\phi^*_{-\sigma},\end{aligned}$$

so that

$$(\phi^*_{-\rho})' = -\tilde{\rho}\sum_\sigma (-\tilde{\sigma})\,R_{\sigma\rho}\,\phi^*_{-\sigma}$$

or

$$(-\tilde{\rho}\phi^*_{-\rho})' = \sum_\sigma R_{\sigma\rho}(-\tilde{\sigma}\,\phi^*_{-\sigma}),$$

i e

$$\psi'_\rho = \sum_\sigma R_{\sigma\rho}\,\psi_\sigma,$$

so that the $\psi_\rho = -\tilde{\rho}\,\phi^*_{-\rho}$ do transform according to the representation Φ One verifies then easily that

$$\psi'_\rho\,\phi'_\sigma - \psi'_\sigma\,\phi'_\rho = \sum_\alpha\sum_\beta \tfrac{1}{2}(R_{\rho\alpha}R_{\sigma\beta} - R_{\rho\beta}R_{\sigma\alpha})\,(\psi_\rho\,\phi_\sigma - \psi_\sigma\,\phi_\rho),$$

so that the spur of the representation subtended by the products

$$\psi_\rho \phi_\sigma - \psi_\sigma \phi_\rho$$

is given by

$$\sum_\alpha \sum_\beta \tfrac{1}{2}(R_{\alpha\alpha} R_{\beta\beta} - R_{\alpha\beta} R_{\beta\alpha}) = \tfrac{1}{2}(\sum_\alpha R_{\alpha\alpha} \sum_\beta R_{\beta\beta} - \sum_\alpha R^2_{\alpha\alpha})$$

$$= \tfrac{1}{2}\{\chi^2(R) - \chi(R^2)\}$$

This is the spur of the antisymmetrical product representation $\{\Phi^2\}$

We have thus shown that the integrals $\int \psi_\rho \phi_\sigma V_r \, d\tau$ and consequently the integrals $\int \phi^*_\rho \phi_\sigma V_r \, d\tau$ transform according to the representation $V\{\Phi^2\}$, for an odd number of electrons with spin

3—PROOF OF GENERAL THEOREM FOR MOLECULES WITH SPIN

We have seen that if the molecule contains an even number of electrons with spin the results obtained in Part I apply without modification Thus we need consider only molecules containing an odd number of electrons For these the electronic wave functions with spin transform according to two-valued irreducible representations of the group of symmetry and for the purpose of establishing our theorem we list below in Table I the two-valued irreducible representations of all the point groups The two-valued

TABLE I—TWO-VALUED IRREDUCIBLE REPRESENTATIONS OF THE POINT GROUPS

D'_∞	E	R	$2C(\phi)$	C_2
$C'_{\infty v}$	E	R	$2C(\phi)$	σ_v
$E'_{\frac{1}{2}}$	2	-2	$2\cos\frac{\phi}{2}$	0
$E'_{\frac{3}{2}}$	2	-2	$2\cos\frac{3\phi}{2}$	0
$E'_{\frac{2l+1}{2}}$	2	-2	$2\cos\frac{2l+1}{2}\phi$	0

$(0 < \phi < 4\pi)$

TABLE I—(*continued*)

D'_{2p+1}	E	R	$2C$	$2C^p$	$2C^{p+1}$	$2C^{2p}$	$2pC_2$	$2pC'_2$
$C'_{2p+1\ v}$	E	R	$2C$	$2C^p$	$2C^{p+1}$	$2C^{2p}$	$2p\sigma_v$	$2p\sigma'_v$
B'_1	1	-1	-1	$(-1)^p$	1	$(-1)^{p+1}$	i	$-i$
B'_2	1	-1	-1	$(-1)^p$	1	$(-1)^{p+1}$	$-i$	i
E'_1	2	-2	$-2\cos\omega$	$(-1)^p 2\cos p\omega$	$2\cos\omega$	$(-1)^{p+1}2\cos p\omega$	0	0
E'_2	2	-2	$-2\cos 2\omega$	$(-1)^p 2\cos 2p\omega$	$2\cos 2\omega$	$(-1)^{p+1}2\cos 2p\omega$	0	0
E'_p	2	-2	$-2\cos p\omega$	$(-1)^p 2\cos p^2\omega$	$2\cos p\omega$	$(-1)^{p+1}2\cos p^2\omega$	0	0

$$\omega = \frac{2\pi}{2p+1}$$

B'_1 and B'_2 are conjugate complex, taken together they give

E'_0	2	-2	-2	$2(-1)^p$	2	$2(-1)^{p+1}$	0	0

D'_{2p}	E	R	$2C$	$2C^2$	C^p	$2C^{p+1}$	pC_2	pC'_2
$C'_{2p,v}$	E	R	$2C$	$2C^2$	C^p	$2C^{p+1}$	$p\sigma_v$	$p\sigma'_v$
$S'_{4p\ v}$	E	R	$2S$	$2S^2$	S^p	$2S^{p+1}$	$p\sigma_v$	pC_2
$E'_{\frac{1}{2}}$	2	-2	$2\cos\frac{\omega}{2}$	$2\cos\omega$	0	$2\ 2\cos\frac{\omega}{2}$	0	0
$E'_{\frac{3}{2}}$	2	-2	$2\cos\frac{3\omega}{2}$	$2\cos 3\omega$	0	$-2\cos\frac{3\omega}{2}$	0	0
$E'_{\frac{2p-1}{2}}$	2	-2	$2\cos\frac{2p-1}{2}\omega$	$2\cos(2p-1)\omega$	0	$-2\cos\frac{2p-1}{2}\omega$	0	0

$$\omega = \frac{2\pi}{2p}$$

T'	E	R	$4C_3$	$4C_3^2$	$4C_3^4$	$4C_3^5$	$6C_2$
E'_1	2	-2	1	-1	-1	1	0
E'_2	2	-2	ϵ^*	$-\epsilon$	$-\epsilon^*$	ϵ	0
E'_3	2	-2	ϵ	$-\epsilon^*$	$-\epsilon$	ϵ^*	0

$$\epsilon = e^{\frac{2\pi i}{3}}$$

TABLE I—(*continued*)

E'_2 and E'_3 are conjugate complex, taken together they give

G'	4	-4	-1	$+1$	$+1$	-1	0

O^r T^r_d	E E	R R	$8C_3$ $8C_3$	$8C_3^2$ $8C_3^2$	$6C_4^2$ $6S_4^2$	$12C_2$ 12σ	$6C_4$ $6S_4$	$6C_4^3$ $6S_4^3$
E'_1	2	-2	1	-1	0	0	$\sqrt{2}$	$-\sqrt{2}$
E'_2	2	-2	1	-1	0	0	$-\sqrt{2}$	$\sqrt{2}$
G'	4	-4	-1	1	0	0	0	0

I^r	E	R	$12C_5$	$12C_5^2$	$12C_5^3$	$12C_5^4$	$20C_3$	$20C_3^2$	$30C_2$
E'_1	2	-2	$\frac{1-\sqrt{5}}{2}$	$\frac{-1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$	$\frac{-1+\sqrt{5}}{2}$	1	-1	0
E'_3	2	-2	$\frac{1+\sqrt{5}}{2}$	$\frac{-1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$	$\frac{-1-\sqrt{5}}{2}$	1	-1	0
G'	4	-4	1	-1	1	-1	-1	1	0
I'	6	-6	-1	1	-1	1	0	0	0

irreducible representations of the crystallographic groups D_2, D_4, D_6 and O have been tabulated by Bethe (1929), and we have followed his method in calculating the characters of the representations of the general groups D_{2p}, D_{2p+1} The characters of the two-valued irreducible representations of the icosahedral group I (and of the groups T and O) have been given already by Frobenius (1899) For completeness we include also the axial groups D_∞, $C_{\infty v}$, but following Tisza (1933) and Bethe we do not include the groups which are the direct product of the inversion or a reflexion with a group already listed, since the representations of the new group can be written down at once In accordance with Bethe's method we introduce a new symmetry R commuting with all the symmetry elements and which denotes a complete rotation through 2π about any axis of the molecule This element is added to the original symmetry group to form the new "double" group which we designate by an index r added to the symbol of the simple group It is to be noted that the double group is not simply the direct product of R with the original group since the relations between the elements of the latter are also altered (e g in D^r_{2p} we have $C_2^2 = R$, whilst in D_{2p} $C_2^2 = E$, C_2

denoting a rotation through π about one of the twofold axes) The two-valued representations of the symmetry group are then one-valued representations of the double symmetry group and their characters can be calculated in the usual way

In order to establish the theorem we have to investigate whether a molecule of a given symmetry possesses at least one set of non-totally symmetrical normal displacements transforming according to a representation V such that $V\{\Phi^2\}$ contains the identical representation, Φ being any two-valued representation of the corresponding group of symmetry V is excluded from being the identical representation, because, as before, the molecular configuration is always taken to be stable with respect to all totally symmetrical displacements Now it is easy to show (see below) that the antisymmetrical product of any two-dimensional representation is the unit representation, so that $V\{\Phi^2\}$ can never contain the identical representation for any two-dimensional representation when V itself is not the unit representation Thus we need list the antisymmetrical product of only those two-valued representations which are more than two-dimensional and these are listed in Table II

TABLE II

Group	Antisymmetrical product of more than twofold degenerate two valued representations
T	$\{G'^2\} = A + E + F$
T_h	$\{G_g'^2\} = \{G_u'^2\} = A_g + E_g + F_g$
T_d and O	$\{G'^2\} = A_1 + E + F_2$
O_h	$\{G_g'^2\} = \{G_u'^2\} = A_{1g} + E_g + F_{2g}$
I	$\{G^2\} = A + H$ $\{I'^2\} = A + G + 2H$
I_h	$\{G_g'^2\} = \{G_u'^2\} = A_g + H_g$ $\{I_g'^2\} = \{I_u'^2\} = A_g + G_g + 2H_g$

The above mentioned property of the two-dimensional representations is established as follows The character of any group element R in the antisymmetrical product is given by

$$\{\chi^2\}(R) = \tfrac{1}{2}(\chi^2(R) - \chi(R^2)),$$

and when R is a two-dimensional matrix

$$R = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

then
$$R^2 = \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & (a_{11}+a_{22})\,a_{12} \\ (a_{11}+a_{22})\,a_{21} & a_{12}a_{21}+a_{22}^2 \end{pmatrix}$$

Thus
$$\chi(R) = a_{11}+a_{22},$$
$$\chi^2(R) = a_{11}^2 + a_{22}^2 + 2a_{11}a_{22},$$
$$\chi(R^2) = a_{11}^2 + a_{22}^2 + 2a_{12}a_{21}$$

Consequently
$$\{\chi^2\}(R) = a_{11}a_{22} - a_{12}a_{21} = \mathrm{Det}(R)$$

Now the determinant of the two-dimensional matrices is always + 1, since they can be expressed as the product of the matrix of a pure rotation with that of the inversion, the determinant of both of which is + 1, the inversion being the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Thus a twofold degeneracy cannot produce instability for a molecule containing an odd number of electrons with spin

Making use of the list of normal displacements of all possible symmetrical molecules given in Table I of Part 1 it is easy to verify that the general theorem is true as formulated above Thus, for example, the antisymmetrical product of the more than twofold degenerate two-valued representations G', I' of the group I always contain the representation H and molecules of the symmetry I always possess at least one set of normal displacements transforming according to H Hence $H\{G'^2\}$, $H\{I'^2\}$ always contain the identical representation, and the symmetrical molecular configuration must be unstable in any more than twofold degenerate electronic state, if the molecule contains an odd number of electrons with spin

4—MAGNITUDE OF THE SPIN EFFECTS*

In considering quantitatively the effects arising from the spin we shall have to compare the order of magnitude of the changes produced by the nuclear displacements in the electrostatic interaction energy and in the spin orbit interaction (multiplett splitting) respectively

Consider first a non-degenerate orbital state in which a certain symmetrical nuclear configuration is stable when only the electrostatic interaction is considered The electrostatic energy in its dependence upon any one normal coordinate d may then be approximated by a parabola

---

* The considerations of this paragraph are due to Dr E Teller

$$\epsilon = \alpha d^2$$

The constant α may be estimated from the binding energy ϵ_0 if r denotes a distance of the order of the equilibrium nuclear distances we will have

$$\epsilon_0 = \alpha r^2 \quad \text{or} \quad \alpha = \frac{\epsilon_0}{r^2}$$

Suppose now that this non-degenerate orbital state has spin degeneracy For the symmetrical equilibrium configuration the spin-orbit interaction will produce a splitting into the various levels of a multiplett Since the orbital state is non-degenerate and possesses therefore no magnetic moment this multiplett splitting will be small, of the order of

$$\epsilon_1 = \frac{I^2}{\Delta E},$$

where I denotes the spin-orbit interaction operator and ΔE the energy difference between orbital states for which it has non-vanishing matrix elements Because of the symmetry the lowest level of this multiplett may still be degenerate Suppose this degeneracy is more than twofold then according to our theorem for small nuclear displacements there will be a linear splitting of the ground state The question is whether this dependence of the spin-orbit interaction upon nuclear displacements can cause any appreciable instability of the original equilibrium configuration To judge of this we must compare its order of magnitude with that of the electrostatic perturbation described by the parabola above The spin-orbit interaction in its dependence upon the normal co-ordinate d may be approximated by the linear relation

$$\epsilon = \beta d$$

The constant β may again be estimated by considering a distance r of the order of the atomic equilibrium distances The process of bringing the atoms together will not change the spin-orbit interaction by an amount greater *than the total multiplett splitting and we may therefore put*

$$\epsilon_1 = \beta r \quad \text{or} \quad \beta = \frac{\epsilon_1}{r}$$

We find therefore for the electrostatic perturbation

$$\epsilon = \frac{\epsilon_0}{r^2} d^2$$

and for the spin-orbit perturbation

$$\epsilon = \frac{\epsilon_1}{r} d ,$$

and we see that whilst for small distances d the latter perturbation is greater than the first, at large distances the reverse is the case Thus each of the energy levels into which the multiplett levels are split may be approximated for large distances by parabolas which will have their minimum slightly displaced from the original minimum (see fig 1) By equating the two perturbation energies we find a distance d which will give a rough estimate of the change in the equilibrium distances which can be caused by the spin-orbit interaction We find

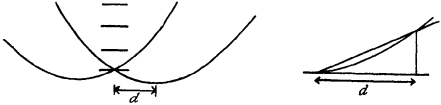


Fig 1

$$d = \frac{\epsilon_1}{\epsilon_0} r ,$$

which is a very small fraction of the equilibrium distance If we take the binding energy ϵ_0 as $1\,\mathrm{V} = 8000\,\mathrm{cm}^{-1}$ and the multiplett splitting ϵ_1 as $1\,\mathrm{cm}^{-1}$, the distance d will have the order of 10^{-4} A This shows that the change of configuration involved is in general so small that it will be covered even by the zero point amplitude of the nuclear vibrations Thus for all practical purposes the spin in this case produces no instability

There is the further possibility that although the lowest state of the multiplett is stable and does not split yet its energy level for small nuclear displacements should intersect the perturbed energy level of a higher state of the multiplett thereby becoming unstable It follows however at once from the above considerations that such an intersection of the parabolas from different levels of the multiplett is impossible Thus in no case can the spin forces produce instability of an orbitally stable non-degenerate state

Let us consider now a degenerate orbital state and ask whether a nuclear configuration which is unstable when the orbital motion alone is considered can be rendered stable when the spin is introduced The multiplett splitting is here greater, being proportional to the spin-orbit interaction operator I

Suppose now that the lowest state of the multiplett is stable for the symmetrical configuration in question, being either non-degenerate or only two-fold degenerate Because of the orbital instability the energy of some of the higher states of the multiplett will show a linear dependence upon the nuclear displacement d

$$\epsilon = \beta d$$

For large displacements r this perturbation will be of the order of the binding energy ϵ_0 (energy between different multipletts)

$$\epsilon_0 = \beta r$$

and we may easily estimate the distance d at which this linearly perturbed energy level will intersect the parabola of the lowest state of the multiplett We find (see fig 2) that this distance is less than

$$d = \frac{\epsilon_1}{\epsilon_0} r,$$

and is thus again a small fraction of the equilibrium distances (we may take $\epsilon_1 = 100$ cm $^{-1}$), being, however, greater than the corresponding distance discussed for the non degenerate state Thus a symmetrical nuclear configuration might under exceptional circumstances be rendered stable in a degenerate orbital electronic state, if the spin orbit interaction is large of the order of the binding energy, but it is not possible for the spin-orbit interaction to cause instability of an orbitally stable state

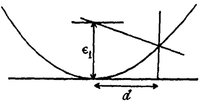


FIG 2

5—CONCLUSION

In conclusion, let us discuss briefly the applicability of our theorem to crystals Bethe (1929) has discussed the splitting of the degenerate states of an atom when placed in the symmetrical field of a crystal and had discussed also the further splitting which occurs when the crystal symmetry is reduced The question could also be discussed by treating the whole crystal as a single molecule and applying our theorem The question then arises why the orbitally degenerate inner shells which occur in the paramagnetic crystals, e g in the rare earth ionic salts, do not cause instability of the crystal lattice Two reasons may be given for this First, we may show by quite similar reasoning to that used above that the linear splitting of the levels of the inner shells is so much smaller than the binding energy of the crystal which arises from the electrostatic interaction of the outer electrons, that the change of equilibrium configuration involved is negligibly small Secondly, we should note that this linear splitting of the inner levels has the same order of magnitude as the perturbation arising from the exchange of these inner electrons between different atoms of the crystal, i e from the possibility of the electrons being propagated through the crystal These translational effects have been treated neither by Bethe nor by us in our theorem Their treatment would involve a consideration of the whole space group of the crystal, whereas Bethe and we ourselves have explicitly restricted ourselves to those groups of symmetry which leave one point of the system invariant

The author is indebted to Dr E Teller for very helpful discussion, and to the Managers of the Royal Institution for facilities granted to complete this research in the Davy-Faraday Laboratory

SUMMARY

It is shown that a polyatomic molecule cannot possess a stable non-linear nuclear configuration in an electronic state having spin degeneracy unless this degeneracy is the special twofold one which can occur only when the molecule contains an odd number of electrons Instability caused by the spin alone is shown to be of secondary importance compared with the orbital effects discussed in a previous paper Table I gives the irreducible two valued representations of all the point groups and will be useful in discussion of the electronic states of polyatomic molecules containing an odd number of electrons with spin

REFERENCES

Bethe, H A 1929 *Ann Phys*, *Lpz*, **3**, 133
Frobenius, G 1899 *S B preuss Akad Wiss* p 339
Frobenius, G and Schur, 1906 *S B preuss Akad Wiss*, p 186
Jahn, H A and E Teller 1937 *Proc Roy Soc* A, **161**, 220
Kramers, H A 1930 *Proc Acad Sci Amst* **33**, 959
Tisza, L 1933 *Z Phys* **82**, 48
Wigner, E 1932 *Nachr Ges Wiss Gottingen*, p 546

Progressive Lightning
IV—The Discharge Mechanism

By B F J Schonland, M A , Ph D

The Bernard Price Institute of Geophysical Research,
University of the Witwatersrand

(*Communicated by Sir Charles Boys, F R S —Received* 24 *September* 1937)

In three previous papers (Schonland and Collens 1934, Schonland, Malan and Collens 1935, Malan and Collens 1937) in this series the writer and his collaborators have described the results of a study of the lightning discharge to ground by means of the Boys and other cameras During the last two years these studies have been continued with improved instruments, and the electrical changes taking place during the discharge have been examined by means of a cathode-ray oscillograph An account of some important results obtained by the latter method has recently been published by Appleton and Chapman (1937) These are confirmed and extended by our own later investigations

With this fuller information it is now possible to put forward some deductions as to the discharge mechanism

1—THE POLARITY OF INDIVIDUAL STROKES TO GROUND

In all cases so far examined by the photographic method the lightning discharge between a thundercloud and the ground starts from the cloud as a leader streamer travelling downwards The circumstances governing the propagation of such a streamer are known from gas-discharge theory to depend upon whether it starts from a cathode or an anode It is therefore necessary at the outset of any discussion of the discharge mechanism to determine the polarity of the cloud end of a lightning stroke

As a result of numerous observations in many parts of the world it is now generally accepted that the total charge conveyed to ground in the great majority of discharges is negative This would seem to indicate that the polarity of the discharge is in general such as to make the cloud end a cathode It has, however, never been established that this conclusion holds in detail for each of the component strokes which make up a discharge Indeed Norinder (1936) has reported that observations made with a cathode-ray oscillograph in Sweden show that a positive discharge stroke is sometimes followed by a number of negative ones

For the South African lightning discharges which have been studied photographically, however, the conclusion suggested by previous studies of the nett change of field (Schonland 1928, Halliday 1932), that the discharge proceeds from a cloud cathode, has now been established for each separate stroke of a composite flash to ground Seventy such flashes comprising some 300 separate strokes to ground have been examined with a cathode-ray oscillograph whose resolving power extends to 50 μsec All the strokes of each flash were recorded photographically and in each case the field changes pointed unambiguously to the slow lowering and sudden destruction of a negative cloud charge

2—GENERAL SURVEY OF THE DISCHARGE PROCESS

Each of the successive strokes or partial discharges which make up a complete lightning discharge to ground takes place in two stages, the downward-moving leader stage being followed by an upward-moving return stage These processes will be described as the leader and the return streamer, since they have the same properties as electrical streamers produced in the laboratory Such a streamer is a conducting filament of ionized gas which extends its length by virtue of ionizing processes occurring in the strong field in front of its tip It is electrically charged throughout its length* but is not at the same potential as the electrode from which it started, for there is a drop of potential along it This drop of potential provides a field which drives a current through the stem of the streamer and this current serves to charge up newly formed sections of the stem to the potential necessary for further progress The current continues at the tip of the streamer as a convection current due to the charge situated there and beyond the tip as a displacement current (§ 7) Light is emitted by the streamer as a result of excitation processes at the tip (§ 8) Apart from this the luminosity associated with the streamer is small and it can therefore be inferred that the field in the stem behind the tip is insufficient to cause excitation by electron impact

It follows from § 1 that all leader streamers observed by us are negative or cathode streamers, and all return streamers are positive or anode streamers The leader process in the case of strokes subsequent to the first is a continuous one and will be termed the dart streamer That for the first stroke proceeds in a series of steps, each step requiring the development of a new step streamer

---

* Cf note p 134

The leader streamer of the first stroke lowers negative charge into the air and distributes it* over the conducting system formed by the leader channel and its branches in the manner shown in fig 1 *a* It was suggested from the photographic studies (Schonland, Malan and Collens 1935, p 622) that this charge represented a considerable fraction of that tapped by the leader This is confirmed by observations of the electrical field change, for Appleton and Chapman (1937) report that the leader or "*a*" stage observed by them causes a considerable change of the thundercloud moment Our own oscillographic records show that the first leader lowers into the air about 85 % of the total charge tapped by it

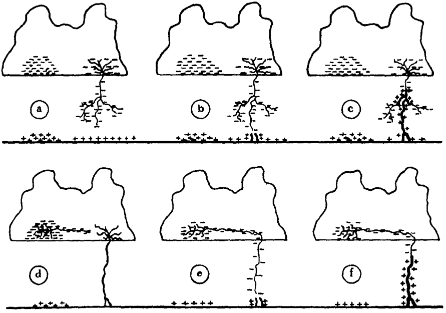


FIG 1

Our observations also show that in the case of strokes after the first the fraction of the charge lowered is much less than in the first leader This is to be explained by the fact that these leaders do not usually have branches The leader system has therefore a smaller capacity and is less effective in lowering charge from the cloud

The second or return stage of the discharge is initiated just before the arrival of one of the leader branches at the ground, by the upward passage of positive streamers from the earth These streamers have been photographed (McEachron and McMorris 1937), and the frequent presence of more than one can be inferred from the root branching shown at the base of many lightning channels (Schonland, Malan and Collens 1935, p 602) This stage is shown in fig 1*b*

---

* These statements, like others in this section, follow what appears to be the accepted and correct view as to streamer development, which it is intended to discuss more fully elsewhere

Once contact is established between leader and earth, the positive return streamer passes rapidly up the leader channel Such a streamer, as was first shown by Simpson (1926), must advance by drawing electrons produced as a result of collision processes in front of its tip, inwards to the tip and passing them down to ground via its conducting stem The relatively immobile positive ions left in front of the tip are then responsible for its rapid extension This stage is shown in fig 1*c* The change of electric moment in this stage of the discharge is large and rapid Its identification by Appleton and Chapman with the second, "*b*", stage of the field change record is supported by our own observations

The last portion of the return stroke process involves the removal and passage down the channel of the residual charge on the cloud centre tapped by the stroke This stage, which is shown photographically by a continuance of channel luminosity after the return streamer has entered the cloud and which is of comparatively long duration, may be identified with part of the final or "*c*" portion of the field-change record

It is implied in this description of the discharge process that separate strokes tap different centres of charge within the same thundercloud, a suggestion based upon evidence discussed in § 10 The general nature of the discharge process for a second stroke from a new charge centre is illustrated in figs 1*d*, *e* and *f*

3—THE LEADER PROCESS

The difference in the behaviour of leaders to first and to subsequent strokes implies a vital difference in the mechanism of their advance The dart streamer invariably follows the path traced out by a previous stroke, even to shifting its track laterally if a wind has blown this path aside (Schonland 1937) This and other features connected with its velocity under different conditions (§ 4) indicate that its mechanism is that of streamer advance along a previously ionized channel, which is further discussed in § 4

The stepped leader, on the other hand, is characteristic of advance into what is apparently virgin, unionized air and involves a different mechanism Associated with it are two velocities, for while the step streamers advance at more than 10^9 cm /sec , the effective velocity of the process as a whole is much lower, most frequently only 1 5 × 10^7 cm /sec It has never been observed to fall below 1 0 × 10^7 cm /sec , though much lower values would not have escaped observation This minimum velocity in the step process is an important clue to the mechanism involved in it, for it can be shown that there is strong presumptive evidence that it is the real velocity of a preliminary streamer which precedes the step streamer

A negative streamer can be imagined to advance into virgin air in one of two ways, as a result of ionization produced either by electrons situated in its tip or by photoelectrons generated in front of its tip In either case a lower limit to streamer velocity is set by the existence of a critical field strength, E_c, in front of the tip, below which ionization by electron impact cannot occur The corresponding critical electron drift velocity, v_c, must be the minimum velocity of advance of a negative streamer, since the streamer as a whole cannot move more slowly than the ionizing agents which produce its extension (Schonland 1935, Goodlet 1937)

It is usual to write the relation between v_c and E_c in the form

$$v_c = \sqrt{(2E_c e\lambda/\pi m)}$$

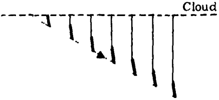
where, following Townsend, the assumption is made that every collision between electron and gas molecule is inelastic Since $E_c \times \lambda$ is known to be sensibly independent of pressure between 54 cm Hg (the mean pressure involved in these discharges) and 76 cm , we substitute $E_c = 30{,}000$ volts/cm and $\lambda = 3\ 8 \times 10^{-5}$ cm , the gas-kinetic electron mean free path at 20° C and 76 cm Hg, and find for the critical velocity, v_c, the value 3 6 × 10^7 cm /sec

Two factors combine to make this result too high In the first place it is known that for nitrogen the value chosen for λ is too large by a factor ranging from 0 8 to 0 3 according to the velocity with which the electrons are moving (Klemperer 1933, p 290) Secondly, the assumption that all collisions are inelastic when the limiting velocity is approached cannot be correct For purely elastic collisions the value of v_c is found to be less than that given above by the factor $\sqrt[4]{(\pi^2 m/4M)}$ (Compton and Langmuir 1930, p 221) or 0 08, m being the mass of the electron and M that of an "air" molecule Applying the first correction, v_c cannot exceed 1 8 × 10^7 cm /sec , while from the second argument it cannot be less than 1 4 × 10^6 cm /sec It is difficult to see how any closer approximation to v_c can be obtained

The observed minimum effective velocity of the stepped leader process, 1 0 × 10^7 cm /sec , will here be identified with v_c, the minimum velocity of advance of an actual negative streamer This identification is supported by the observations made by Allibone and the writer upon negative leaders in the laboratory spark discharge which appear, as Goodlet has recently pointed out (1937), to show a similar minimum velocity of the order of 1.5×10^7 cm./sec.*

The present suggestion is equivalent to the statement that a true negative virgin air streamer travels continuously downwards in front of the step streamer processes with a velocity equal to the effective velocity of these. Upon this, so far unrecorded, streamer the steps are periodically superimposed. It follows that the step streamers, like the dart streamers of subsequent strokes, travel along a previously ionized channel provided by this slower *pilot streamer* which precedes them.

The suggestion finds immediate support in the relation which has been found to exist between the time of pause of a step streamer, t, and the length of the step, l, which is executed after this pause (Schonland, Malan and Collens 1935, p. 616). For each individual case the ratio l/t is equal to the effective velocity at this stage of the process, that is to the velocity of the suggested pilot streamer. If t is long the subsequent step l is long and vice versa. The steps thus retrace and brightly illuminate an ionized channel formed during the pause period and cease when they have caught up with the tip of the pilot streamer. During the following pause, the origin of which is discussed in § 6, the pilot streamer once more forges ahead. Fig. 2 illustrates what would be observed on a camera with a fixed lens and a film moving from right to left if the pilot streamer, shown as a dotted line, could be recorded as well as the step streamers which follow it. An explanation of the small luminosity associated with this pilot and the consequent difficulty of recording it photographically or by the field-change method is given in § 8.



Film → movement

FIG. 2

It is possible in this way to offer a simple explanation of the tortuous nature and branching of the first leader channel. In the majority of cases the effective or pilot velocity of the process is less than $1\ 5 \times 10^7$ cm /sec , not very far from the critical minimum of 10^7 cm /sec below which it cannot progress at all Its direction will therefore be controlled by the structure of the electric field in the neighbourhood of the pilot streamer tip and by variations in the density of local space charge

---

* The existence of a minimum velocity of this order for negative streamers serves as an additional criterion of polarity. The minimum for positive streamers, if it exists, is much lower, for the writer has recently photographed such streamers with a velocity of 1.6×10^6 cm./sec.

It may be noted that a stepped method of development in the case of Lichtenberg figures from a negative point has been inferred by v Hippel (1933, 1934) under certain experimental conditions and has been interpreted by him in terms of a pause during which a pilot streamer (*vorentladung*) moves ahead of the main spark channel A similar mechanism was proposed by Toepler (1926) to account for beaded lightning (*perlschnurblitz*)

So far the discussion has been limited to the case of pilot streamers travelling with the minimum observed velocity Before it is possible to refer to the behaviour of such streamers at higher velocities it is necessary to consider the effect which photoionization will have upon its development in fields greater than the critical value E_c This question is also involved in the mechanism of all streamers travelling along a previously ionized path In the case of the lightning discharge there are three such types of streamer--the dart leader streamer to subsequent strokes, the return streamer to all strokes and, as has been suggested above, the step streamer which follows the pilot streamer in first leaders

4—THE MECHANISM OF STREAMER ADVANCE ALONG A PREVIOUSLY IONIZED TRAIL

It has been pointed out by Cravath and Loeb (1935) that the very high velocities observed for the dart streamer can only be interpreted satisfactorily in terms of a mechanism which involves the pre-existence of free electrons in the region traversed by the streamer These free electrons they considered to be produced by the action of natural agencies such as radioactive and cosmic radiation In the case of the dart streamer, however, as well as of the other two streamers mentioned above, they arise from the previous passage of a powerful ionizing discharge along the trail to be followed by the streamer

The mechanism of Cravath and Loeb is illustrated by fig 3 *AB* is the streamer stem and the shaded portions of the previously ionized trail in front of the tip of the streamer indicate the region over which the electric field exceeds the critical value for impact ionization by electrons In the case of a negative streamer (dart or step) the original electrons in *BC* create others by moving forward but in the case of a positive streamer (return streamer) they do so by moving inwards to the streamer tip As soon as BC has in this way been rendered sufficiently conducting as a result of the formation of electron avalanches, the streamer has effectively extended to C It is thus unnecessary for any electrons to travel the full distance BC and the streamer velocity V can considerably exceed the mean electron drift velocity $\bar{v}$, of the electrons in BC

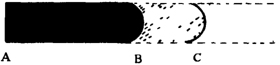


Fig 3

It is here suggested, as an extension of this theory, that consideration must be given to the time necessary for effective ionization of the region BC This is the time required for this region to be effectively covered by a network of small filaments, one starting from each original electron We may assume, as the simplest possible hypothesis, that this is the time, t, necessary for each original electron to cover the mean distance from one electron to the next Then if n is the original electron density in the trail and $\bar{v}$ the mean electron drift velocity in BC, we have

$$t = 1/(n^{\frac{1}{3}} \times \bar{v}) \quad \text{and} \quad V = d/t = n^{\frac{1}{3}} \times \bar{v} \times d, \tag{1}$$

where $d = BC$

The streamer velocity V thus depends on the pre-existing electron density in the ionized trail, n, as well as on the electric field in front of the tip, which determines $\bar{v}$ and d As an illustration we may take the case of a streamer with a tip radius of 1 cm and a tip potential of 3×10^6 volts The capacity of the tip of the streamer may then be estimated at 0 5 e s u and the tip charge at 5×10^3 e s u The field strength immediately in front of the tip is 1 5 × 10^6 volts/cm , and if the streamer were advancing into virgin air its velocity would be about 7×10^7 cm /sec , since the field is fifty times the critical value already discussed If, however, the region into which the streamer was advancing contained 10^3 pre-existing electrons per c c the streamer velocity would be given by equation (1) above In this case the distance d over which ionization took place in front of the tip would be 6 cm , the mean field strength, calculated for a distance 3 cm in front, would be 9×10^4 volts/cm and the mean electron drift velocity 1 7 × 10^7 cm /sec Substituting these values in equation (1) we find that the new velocity of advance V would be 1 03 × 10^9 cm /sec or fifteen times as great as for advance into virgin air The Cravath-Loeb mechanism thus offers a satisfactory explanation of the high velocities of dart and return streamers without the introduction of impossibly high field strengths in front of their tips The similarly high velocity found for the step streamer is another argument in support of the suggestion that this too travels along a previously ionized path

The chief addition which has here been made to the mechanism suggested by Cravath and Loeb is the introduction of the electron density of the trail, n, in equation (1) The streamers we are discussing provide strong evidence that their velocities depend upon n in the manner in which these velocities are affected by the age of the trail they follow, that is by the time which has elapsed since this trail was formed The free electrons produced during the previous ionizing process will tend to disappear by processes of capture and recombination and n will be an inverse function of the age of the trail Thus it is found that fast dart leaders occur when the interval since the preceding stroke is short and slow ones when it is long, and that when the interval is unusually long the continuous dart process cannot occur but is replaced by a rapid stepped proçess (Schonland, Malan and Collens 1935, p 604)

Similar evidence is provided by the return streamer Since this is a positive streamer the electrons in the forward region move inwards, but otherwise equation (1) should apply The very high velocities attained by the return streamer are no doubt partly due to the strong-tip fields caused by the presence of a negatively charged leader channel in front of the positive tip of the streamer The velocity of this streamer is, however, always observed to decrease as it travels up the leader channel, and this fact suggests that the governing factor in its progress is again the age of the trail it follows For the time which has elapsed since the leader was first formed is smallest at the ground end of the channel, where it is zero, and greatest at the cloud end, where it may be 0 02 sec Whether this factor alone is responsible for the decrease observed or whether it is combined with a reduction in tip field due to the reduced negative charge density on the upper portions of the leader channel illustrated in fig 1 it is difficult to determine

As the return streamer travels along a branch, the age of the ionized trail in front of it actually decreases and n must increase Thus in fig 4 the light lettering shows the times in μ-secs at which various parts of the leader channel and its branch $ACFG$ were formed, while the heavy lettering shows the actual age of the ionization at each point at the moment when the return streamer first reaches it At the extreme end of the branch the actual age is little different from that at the base of the channel This diagram shows that the trail electron density n along a branch will be greater than along subsequent parts of the main channel The age at D on the main channel when the return streamer arrives there is 7070 μsec, while that at C, equidistant from A but along the branch, is only 5070 μsec Between E and F, also equidistant from A, there is a difference of 4000 μsec in age When, therefore, the return stroke reaches A it divides into two portions, one of which travels along the branch trail much faster than the other along the main channel The branch streamer will show a higher luminosity

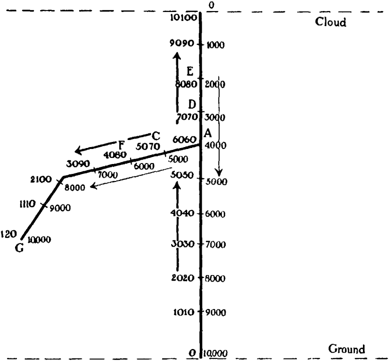


Fig 4

(§ 8) and carry a larger current than the main channel streamer, and the discharge will appear to concentrate its energy upon the branch These effects and the component streamers they evoke at the base of the channel have recently been discussed by Malan and Collens (1937) As in the case of the main channel streamer discussed in the previous paragraph it is possible that the "age" effect just described is combined with an increase in tip field, for the density of the negative charge on the leader will be greater along a branch than along the main trunk

More detailed studies of the structure of the "b" portions of field change records, which are at present being undertaken, should show electrical evidence of these effects

5—The pilot streamer

According to the view put forward in preceding sections, the pilot streamer is the only one of the four types of streamer involved in the lightning discharge which can properly be said to advance into virgin air These pilot streamers have velocities which have been recorded (as effective step process velocities) as ranging from $1\ 0 \times 10^7$ cm /sec to $2\ 0 \times 10^8$ cm /sec If their mechanism was simply that of ionization by their tip electrons, their velocity would be equal to the drift velocity v of these electrons and we would have $V = v = k\sqrt{E}$ where k is a constant discussed in § 3 Since from § 3, $E = E_c = 30{,}000 \times 54/76 = 21{,}000$ volts/cm at the mean pressure prevailing in these discharges, when $V = v_c = 1\ 0 \times 10^7$ cm /sec , the maximum value observed for V, $2\ 0 \times 10^8$ cm /sec , would correspond to a tip field of 8×10^6 volts/cm The question then arises whether at the higher velocities such high tip fields actually exist or whether in this case also the streamer makes use of pre-existing ionization in front of its tip

The suggestion of Loeb and Cravath that such ionization, by which is implied the presence of free electrons arises from natural sources does not appear to be valid The mean life of a free electron is so short and the natural supply so meagre that it is easily shown that the probability of an electron being present at the required moment in the region in front of the tip is negligibly small ($< 10^{-6}$/c c) There remains a later suggestion by Loeb (1936) that streamer advance into virgin air may be facilitated by photoelectric ionization of the air in front of the tip If this process occurred it would be possible for the maximum pilot-streamer velocities quoted above to be reached by the process represented by equation (1) without the high tip field strengths which have been calculated

It will, however, be shown in § 9 that pilot streamer velocities can be separated into two groups The first group, which includes velocities from $1\ 0 \times 10^7$ to $5\ 0 \times 10^7$ cm /sec , contains the majority of the streamers and evidently involves field strengths not much higher than the minimum for ionization The second includes all velocities from $5\ 0 \times 10^7$ to $2\ 0 \times 10^8$ cm /sec and contains a minority of the streamers whose behaviour indicates that they may be travelling in fields of abnormal strength

There is therefore no need to invoke photoelectric ionization to explain even this minority of faster pilot streamers, and the evidence at present suggests that negative streamer advance into virgin air takes place by ionization directly produced by electrons in the tip of the streamer

6—THE MECHANISM OF THE STEPPED LEADER PAUSE

The time of pause between successive steps in the stepped leader process occupies a remarkably narrow range of values for a phenomenon of this kind In 90 % of the many leaders examined it lies between 50 and 90 μsec This makes it difficult to accept any explanation of the pause which does not involve a gas-discharge mechanism of a very fundamental character Thus it is unlikely that the pause can be connected with any property of the cloud charge which is feeding the leader, for this must show wide variations from one flash to another A similar reason leads one to reject any interpretation based upon oscillations in, or surges travelling along, the channel between cloud and the tip of the leader For such an explanation must always give a pause time which increases with the length of the channel, and such an increase is not observed

According to the view expressed in § 3, the bright step streamer ceases at the end of a step because it is a process which requires a previously ionized trail, and at the end of a step it has caught up with the pilot streamer which lays this trail The pilot streamer then forges ahead during the time which elapses before the step streamer can start

The tip of the arrested step streamer is thus the beginning of the stem of the pilot streamer, and the field in the stem in front of this tip is insufficient for ionization by electron impact (§ 2) The pause time is then to be considered as the time necessary for this tip field to approach the critical value for ionization There seem to be two possible views as to the way in which this growth can occur According to the first it is due to the decrease in the electron density in the stem of the pilot streamer This streamer carries a current (§ 2) and electron drift is taking place in its stem to maintain it If, however, the electron density is falling off with time as a result of capture and recombination processes the current can only be kept up by a rise in the field strength in the stem Ultimately this field is great enough for the step streamer to start

It may be, however, that the pause time is the time required for a positive space charge to form up from the positive ions in the pilot stem and gather close enough to the step-streamer tip to give it a strong starting field This second explanation, which was suggested by v Hippel (1933, 1934) in connexion with Lichtenberg figures, regards the pause time as of the same nature as a spark lag It may, however, be urged against it that the gaseous tip of the step streamer is very different in nature to the metal cathode of a spark gap, and the formation of a positive space-charge layer very close to it is difficult to picture

7—PILOT-STREAMER CURRENTS

An estimate of the current carried by a pilot streamer can be obtained from an expression derived by Rudenberg (1930) on the assumption that the tip of the streamer is hemispherical and the tip charge uniformly distributed over this hemisphere, the charge density along the stem being negligible in comparison with that on the tip The current i is then found to be given by

$$i = E\ r\ V/2, \tag{2}$$

where E is the tip field in e s u , r the tip radius and V the streamer velocity A similar expression has been found by Jehle (1933) for the convection current carried by the tip charge

Values of E and V for the pilot streamer have been given above, but no direct information is available as to the radius r of its tip It is certainly less than 10 cm , since measurements of the diameter of the lightning channel after both leader and return stroke have traversed it give radii less than this (Schonland, 1937) The streamer tip will spread laterally as it advances, and we can apply to it the calculations made by Ollendorff (1932) on the spreading of an electron swarm as a result of collisions with gas molecules These give a radius of 1 cm for the swarm after it has travelled 10 m in a field of strength 30,000 volts/cm Thus for the slowest pilot streamers which do not travel further than 10 m before they are caught up by the step streamer, r will be less than 1 cm Substituting in equation (2), $E = 21{,}000$ volts/cm , $r = 1$ cm and $V = 1{\cdot}0 \times 10^7$ cm /sec , we find that pilot streamers travelling with the minimum velocity observed carry a current of the order of 1 amp

The spreading of the tip is less in the larger fields which produce the higher streamer velocities, r varying as $\sqrt{(l/V^2)}$, where l is the distance travelled The distance which a pilot streamer travels before being caught up by the step streamer is Vt, where t is the pause time (Schonland, Malan and Collens 1935, and § 3) Since t is approximately constant this distance is proportional to V, and the radius of the pilot tip at the end of its career varies as $1/V^{\frac{1}{2}}$ In the case of the fastest pilot observed, for which $V = 2 \times 10^8$ cm /sec , the final-tip radius will thus be of the order of $1/\sqrt{20}$ cm or about 2 mm Substituting these values and that for E previously calculated (8×10^6 volts/cm) in equation (2), we find the maximum current carried by a pilot streamer to be 180 amp Such high velocities are rare, and this discussion indicates that the majority of these streamers involve currents of a few amperes only

These currents are very small compared with those carried by the other types of streamer in the discharge The dart leader removes a charge of the order of 1 coulomb from the thundercloud in a time ranging from 10^{-3} to 10^{-4} sec and thus carries from 1000 to 10,000 amp The return streamer removes this same charge to ground in from 50 to 100 μsec which involves a current of the order of 130,000 amp

8—STREAMER LUMINOSITY

The main source of light in the lightning discharge is situated near the tips of the various streamers which comprise the discharge Thus the dart streamer is so called because the light which it emits is largely confined to a length of some 50 m behind its moving tip, the streamer stem behind this being only faintly luminous The same is true of the step streamer whose stem can only be photographed with difficulty In the case of the return streamer the emitting region is longer, but here also the luminosity at the base of the channel has fallen off considerably before the bright tip of the streamer has advanced half of the distance between ground and cloud

This localization of the main source of light implies that it arises from excitation processes in the strong fields at the tip of the streamer where atoms and molecules are raised to excited states by electron impact Since the tip has passed on before they return to normal states with the emission of light, the luminosity thus produced will be spread over a distance D behind the tip whose length is given by $D \geqq v\tau$, where v is the streamer velocity and τ the mean lifetime of an excited state The sign $>$ indicates the effect of those excited atoms which have a longer lifetime than the mean in lengthening the emitting region

For a dart leader of velocity 10^9 cm /sec , D is observed to be about 50 m and for a return streamer of velocity 10^{10} cm /sec to be about 500 m (Schonland and Collens 1934, pp 662, 665) Both these streamers thus give $\tau \leqq 5 \times 10^{-6}$ sec , for the mean lifetime of an excited state in air at ordinary pressure It is interesting to note that values for τ in the arc spectrum of nitrogen are known to be of the order of 10^{-7} sec (Compton and Langmuir 1930, p 132) It may be added that the observed length, D, of the emitting region is not affected to any important extent by the resolving power of the camera lens used The image of a point on the channel represents the combined effect of the light received from a length l on either side of it, but for a discharge at a distance of 5 km l is only 10 cm

All these streamers show a rapid increase in the brightness of the tip region with increasing velocity (Schonland, Malan and Collens 1935, p 606) This is to be explained by the fact that the velocity of the streamer, according to equation (1), is a function of the mean drift velocity, $\bar{v}$, of electrons in the region ahead of the tip, since the probability of excitation is well known to rise sharply with electron velocity in the neighbourhood of the ionization potential

We may thus infer that the brightness of a very fast streamer implies a high field in front of its tip, a large value of $\bar{v}$ and a degree of ionization which is larger than usual This increased ionization provides the stem of the streamer with the conductivity necessary for it to carry the streamer current without an excessive fall of potential per unit length As a result the tip potential and field can be maintained at the value necessary for its rapid progress

In the case of the pilot streamer the electron-drift velocity at its tip is identical with the streamer velocity V, and this is usually not far from the critical value for ionization The brightness of such a streamer would therefore in general be small and the conductivity of its stem be low These effects would account for the fact that it has not been recorded photographically and for the small currents which it carries It should, however, be possible to observe it at the higher velocities, where the increased brightness of the following step streamer (Schonland, Malan and Collens 1935, p 615) indicates that a trail of greater electron density than usual has been formed

9—LEADER DEVELOPMENT AND SPACE CHARGE

It is now possible, from the photographic evidence available, to distinguish two different types of stepped leader development The most frequent, which may be termed type α, is that in which the leader process moves in a fairly regular manner from cloud to ground without excessive branching In such a leader the pilot or effective velocity is not much greater than the critical minimum for progress and ranges from $1\ 0 \times 10^7$ to $5\ 0 \times 10^7$ cm /sec The field in front of the pilot is therefore not much greater than the critical breakdown field Such leaders we find to be associated with the type of field-change record recently described by Appleton and Chapman and referred to in § 2

The second or β type of leader, which occurs in about 30 % of the cases examined by us, is very much faster and brighter in its initial stages than the α type All the high effective velocities which have been observed are found during the first portion of the path of leaders of this type This portion is always associated with an extensive branching process at its end, and in a number of cases (e g flash 36, fig 9, Schonland, Malan and Collens 1935) the discharge ends in this branched manner in the air More usually, however, the process continues onwards to the ground with reduced velocity and with steps which are faint and short Occasionally the extensive branching and subsequent slower downward development is repeated several times

It is natural to associate this behaviour of the leader process with the existence in the air of concentrations of positive space charge formed by the process of point discharge first suggested as important by Wilson (1925) Such concentrations would lead to the formation of stronger electric fields than usual between the cloud charge and the space charge and explain the higher pilot velocities of the streamers which travel to them Below such a concentration the field would be low and the pilot velocity would be much reduced in the manner mentioned above

Our field-change records show precisely similar effects, the β type leader beginning with a large and rapid field change carrying abnormally large superimposed pulses due to the steps and ending with a small field change which takes place slowly and with no observable step effects

An interesting complication of the type β leader which is sometimes observed is the occurrence of a bright dart streamer along the previously formed section of the channel from cloud to space-charge branch before the further downward movement takes place In a complex leader as many as four of these dart processes can take place at different stages of its progress A moving camera which was unable to record the faint and slow stepped process which preceded the later of these dart streamers would indicate that the discharge took place in four very large steps This type of development has been reported by Walter (1910 1935) and by Workman, Beams and Snoddy (1936)

10—THE ORIGIN OF SEPARATE STROKES

The number of strokes making up a complete discharge to ground is a variable quantity It has been shown (Schonland, Malan and Collens 1935, p 599) that single stroke discharges are the most frequent and that discharges with more than six strokes are rare We have observed that the multiple-stroke discharge is more usually associated with the extensive and violent frontal type of thunderstorm than with the type which owes its origin to thermal convection

The origin of multiple strokes has been given a suggested explanation by Simpson (1926) in terms of his well-known theory of the positive discharge, the cloud end of the positive streamer becoming blocked with negative ions This explanation cannot be applied to the negative discharges which have been studied here Ollendorff (1933) has suggested that the discharge is cut off during each stroke by the effect of the potential drop at its earthed end upon the falling characteristic of the channel This explanation, which involves a much simpler conception of the channel than that actually observed, appears to have been disproved by experiments reported by McEachron and McMorris (1937)

The experimental evidence obtained by us suggests that separate strokes arise as a result of the existence of separate generating centres at different places within the cloud That this can actually happen is shown by the not infrequent occurrence of Y-shaped discharge channels, apparently branching upwards, which in the cases examined by us with the Boys camera are due to two separate strokes from different parts of the cloud with a common stem formed by one stroke and utilized by the other The association of multiple strokes with large and extensive cloud masses is in accord with this suggestion, for the larger thunderclouds may be expected to possess several generating centres When one of these, due perhaps to the presence below it of a sufficient positive space charge, is able to discharge to ground the reorganization of the field within the cloud and its concentration in the direction of the top of the conducting channel is likely to start a leader streamer which, like the second Y leader described above, will use the cloud-earth channel formed by the first

On this view the time interval between one stroke and the next should be of the same order as the time required for a stepped leader to travel from a new cloud centre to one which has previously discharged to ground, plus the time taken by a dart streamer to travel down the previously formed channel to earth Owing to the high velocity of the latter its contribution to the total time interval will be small, and we may write the interval T as D/V, where D is the distance between the new and the old cloud centre and V the velocity of a stepped leader The observed values of T are usually greater than 0 01 and less than 0 09 sec, the most frequent time interval being 0 03 sec The most frequent value observed for V in the air outside the cloud is 2×10^7 cm/sec Substitution of this figure and $T = 0\ 03$ sec yields $D = 6$ km for the distance between cloud centres

This calculation has not taken into account the possibility of a delay in the starting of the new stepped leader nor of a reduced velocity for the stepped leader process in the presence of water-drops, both of which would reduce the value found for D We may conclude that this explanation of the time interval indicates that the distance between cloud centres of charge responsible for separate strokes is less than 6 km, a value which appears very reasonable when it is recalled that they arise in the main from clouds of considerable size It is of interest to note also that the diameter of a single charged cloud centre has been estimated by Wilson (1921) to be 1 0 km

This work forms part of the lightning research programme of the South African Institute of Electrical Engineers I have to thank my colleagues on the committee, in particular Mr H Collens, Dr D J Malan and Mr D B Hodges, for permission to refer to results obtained in collaboration with them which are not yet published, and for much helpful discussion

SUMMARY

1—Oscillographic observations indicate that all processes in the discharge to ground observed in South Africa involve a cloud cathode and an earth anode

2—The first lightning stroke appears to involve

a—The development of a pilot streamer, a negative streamer proceeding from the cloud into virgin air

b—The periodical catching up of this pilot streamer by a much faster step streamer, a negative streamer advancing along an ionized path

c—The distribution by this leader process of the greater portion of the cloud charge tapped by it upon a branched leader channel in the air below the cloud

d—The passage of this charge to ground in the return stroke, a positive streamer travelling along an ionized and oppositely charged path

3—The second and subsequent strokes involve

a—A fast dart streamer, a negative streamer advancing along an ionized path

b—A return stroke streamer similar to 2*d*

4—The mechanisms of the three types of streamer are investigated Satisfactory explanations of their behaviour can be derived

5—Discussions are given of the currents in the various streamer processes, of the luminosity associated with their movement and of the effect of space charge on leader development

6—Evidence is given which indicates that the occurrence of separate strokes in the discharge is due to the presence in the cloud of separate charge-generating centres

References

Appleton, E V and Chapman, F 1937 *Proc Roy Soc* A, **158**, 1
Compton, K and Langmuir, I 1930 *Rev Mod Phys* **2**, No 2
Cravath, A and Loeb, L B 1935 *Physics*, **6**, 125
Goodlet, B L 1937 *J Instn elect Engrs* (in the Press)
Halliday, E C 1932 *Proc Roy Soc* A, **138**, 205
Hippel, A v 1933 *Z Phys* **80**, **19**
— 1934 *Naturwissenschaften*, **22**, 47
Jehle, H 1933 *Z Phys* **82**, 784
Klemperer, O 1933 "Electronik", p 290 Berlin Springer
Loeb, L B 1936 *Rev Mod Phys* **8** 267
McEachron, K B and McMorris, W 1937 *Gen Elect Rev* **39**, 487
Malan, D J and Collens, H 1937 *Proc Roy Soc* A, **162**, 175
Norinder, H 1936 *J Franklin Inst* **225**, 69
Ollendorff, F 1932 *Arch Elektrotech* **26**, 1932
— 1933 *Arch Elektrotech* **27**, 169
Rudenberg, R 1930 *Wiss Veröff Siemens Konz* **9**, part 1
Schonland, B F J 1928 *Proc Roy Soc* A, **118**, 233
— 1935 *Nature, Lond*, **136**, 1039
— 1937 *Phil Mag* **23**, 503
Schonland, B F J and Collens, H 1934 *Proc Roy Soc* A, **143**, 654
Schonland, B F J, Malan, D J and Collens H 1935 *Proc Roy Soc* A, **152**, 5
Simpson, G C 1926 *Proc Roy Soc* A, **111**, 56
Toepler, M 1926 *Mitt Hermsd Schomb Isolat G m b H* **25**, 743
Walter, B 1910 *Jb hamburg wiss Anst* **27**
— 1935 *Ann Phys, Lpz*, **22**, 421
Wilson, C T R 1921 *Philos Trans* A, **221**, 73
— 1925 *Proc Phys Soc* **37**, 32*d*
Workman, E J, Beams, J W and Snoddy, L B 1936 *Physics*, **7**, 375

The Theory of the Photolysis of Silver Bromide and the Photographic Latent Image

By R W Gurney and N F Mott, F R S

H H Wills Physical Laboratory, University of Bristol

(*Received* 21 *October* 1937)

1—Introduction

The aim of this paper is to see to what extent the processes taking place during the exposure and development of a photographic emulsion can be understood in terms of the conceptions of present day quantum mechanics The problem has recently been discussed by Webb (1936), to whose paper we are considerably indebted, we believe it possible, however, to push the analysis further than has been done by him

We may conveniently divide our discussion into two main parts, first, that dealing with the direct photolytic reduction of silver halides, and second, that dealing with the latent image and its subsequent chemical or physical development

PART I—PRINT OUT EFFECT

2—Direct photolytic reduction of silver halides

We must first summarize what is known about this process from the experimental side Hilsch and Pohl (1930) have investigated the photochemical behaviour of large crystals of silver halides They find that under illumination the crystals become coloured, a new absorption band appearing in the visible They attribute this absorption band to the formation of colloidal particles of metallic silver * A photoelectric current may be measured during the photochemical coloration (cf Lehfeldt 1935) The quantum efficiency of the process is not known The effect of temperature has been investigated by Lohle (1933), who finds that at – 186° C no colloidal silver is formed unless the crystal has previously undergone a brief illumination at room temperature

In photographic emulsions the quantum efficiency of the formation of silver has been investigated by Eggert and Noddack (1923) and others, and it is established that in the direct reduction of silver-bromide emulsions by light (print-out effect), approximately one atom of silver is formed for every quantum absorbed This only holds for room temperature, at liquid air temperature Berg and Mendelssohn (1937) have shown that the efficiency falls by 10^5

---

* These must be sharply distinguished from the centres of atomic size, which give rise to the *F* band in *alkali* halides

It must be emphasized that in a partially printed out emulsion also the silver is not found in atomic form, but, in each halide grain, in a relatively small number of specks of colloidal metal, as may be seen by examining an exposed grain under the microscope (cf Hay 1932*a*) The light quanta incident must, however, be absorbed at random all over the grain, the silver is thus not formed *where* the light quantum is absorbed It is often stated that "silver atoms are formed where the light is absorbed, and that the atoms then flock together to form colloidal metal" In what follows we hope to give a more adequate description of this process

3—PHOTOCONDUCTIVITY

As we have stated the absorption of light by silver halides is accompanied by photoconductivity Now it is a well-known result of the quantum theory of solids that if an electron is brought from outside and placed on one of the metal ions of an alkali or silver halide, it will not stay on that ion, but will be able to jump from ion to ion through the crystal Unpublished calculations for NaCl by Fuchs and the present authors suggest that the potential barrier preventing this interchange is practically negligible, such an electron can run through the crystal as quickly as a free electron in a metal We refer to the band of energy levels that such an electron can have as the "conduction levels" Photoconductivity is to be ascribed to the raising by the light of an electron into the conduction level from a bound state in one of the halogen ions At present we shall assume that every quantum of light absorbed does this—an assumption which is discussed in greater detail in Part III of this paper

We may estimate the diffusion coefficient of an electron in the conduction band If it has thermal energies its velocity will be $\sim 10^7$ cm /sec , and its mean free path,* say 10^{-7} cm , so its diffusion coefficient is of the order 1 cm ^{2}/sec

If then we consider a grain of a halide emulsion, of linear dimensions, say 10^{-4} cm , absorbing quanta at the rate 100/sec , we see that the electrons will diffuse rapidly away from the points where they are released, and form a sort of *electron gas* in the grain

---

* Much less than that of an electron in a metal (cf Frohlich 1937)

It is not necessary to assume that these electrons are actually moving about all the time According to Pohl (1937), if a crystal of AgBr is irradiated at low temperatures with 10^{12} quanta/cm ^{3}, it shows electronic conductivity like a semi-conductor The state is destroyed by high temperatures or by high fields We may assume then that there exist in the crystal metastable positions at cracks or impurities at which the electron may temporarily be trapped, and from which it will be released later by thermal motion This follows also from the fact that in Pohl's experiments on photoconductivity a field of about 1000 V/cm was necessary for the current to rise to its saturation value

Now consider what happens in a grain already partially reduced, so that it already contains some metallic silver Whenever an electron comes in contact with a speck of silver it will be captured This is shown in fig 1, where the energy levels of an electron in the halide crystal and in metallic silver are shown, the electron will fall down the potential hill into the metal and not be able to get out *

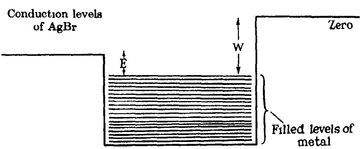


FIG 1—Conduction levels of AgBr in contact with metallic silver, W is the work function of silver

It is possible that a few electrons will be captured by some of the halogen atoms from which they have been ejected If, however, there is already a considerable quantity of metallic silver present, this will be improbable, since there will at any moment be very few of these atoms (cf § 6) Thus eventually an electron will find its way to the metallic silver, and we may assume that for every quantum absorbed an electron is added to the metallic silver

---

* There is experimental evidence that electrons moving in the conduction levels of salts are easily trapped by colloidal particles of metal; Gläser and Lehfeldt (1936) find that the mean distance moved by photoelectrically released electrons in KCl crystals containing F centres is considerably diminished if the centres are allowed during heat treatment to form colloidal particles of metal Lehfeldt (1935) reports the same effect in silver salts

4—ELECTROLYTIC CONDUCTIVITY

Silver halides are electrolytic conductors in the solid state Below room temperature the conductivity is structure sensitive (cf Lehfeldt 1933) Above room temperature it is not, and shows a temperature dependence of the type

$$\sigma = Ae^{-B/T}$$

The following mechanism has been proposed for the electrolytic conduction (cf Jost 1937) In thermal equilibrium at temperature T a few of the positive ions will have left their normal positions and will be found in the centres of cubes at large distances away, as shown in fig 2 We call such positions "interlattice" positions If E_0 is the work necessary to remove an ion to an "interlattice" position in this way, the fraction of the total number so displaced will be

$$e^{-\frac{1}{2}E_0/kT}$$

Both the displaced ion, and the "hole", will be able to move through the crystal, a certain activation energy W_0 being necessary in either case The conductivity due to either process will then be proportional to a factor of the type

$$e^{-(\frac{1}{2}E_0+W_0)/kT}$$

Both processes are due to the motion of silver ions

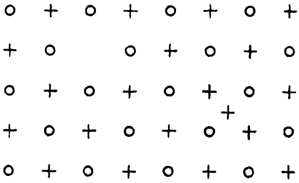


FIG 2—Displaced ion in AgBr lattice

We now return to our halide grain containing negatively charged particles of silver The silver ions in interlattice positions will be attracted to the charged silver,* and will therefore move through the crystal until they meet and adhere to the silver The silver specks will therefore grow until their charge is neutralized We thus see why an atom of silver is formed for every quantum absorbed

When an electron is removed from a halogen ion by the light, the vacant place ("positive hole") left behind may perhaps have a certain mobility, since an electron can move from a neighbouring ion and fill it up, and so on through the lattice Unless this mobility is much less than that of the silver ions in interlattice positions, our whole model breaks down, because the negative charge on the silver will be neutralized by the positive holes, which are attracted to it We shall assume, then, that the positive holes are not mobile, they may be trapped by the mechanism proposed by Gurney and Mott (1937)

In § 10 it is suggested that the light is absorbed mainly on the surface of the halide grains, if that is the case there is no difficulty in understanding how the bromine *atoms* so formed escape into the gelatine

5—PRINT-OUT EFFECT AT LOW TEMPERATURES

As we have seen, the quantum efficiency of the print-out effect, near to unity at room temperatures, falls by a factor 10^5 as the temperature is lowered to that of liquid air (Berg and Mendelssohn 1937) This may be due to two causes *either* the production of electrons may be less efficient at low temperatures, *or* electrolytic conductivity may take place only very slowly In the latter case the silver present would charge up negatively, but it could not rapidly discharge by attracting silver ions to it, so electrons subsequently formed would be repelled and would wander round the grain until recombination took place with a halogen atom

If the former explanation is true we should expect silver bromide to show a marked drop in the photocurrent per absorbed quantum as the temperature is lowered Such a drop is observed in zinc-blende and in coloured alkali halides and is to be expected on theoretical grounds (cf Mott 1937), though possibly at a lower temperature than that of liquid air In AgBr Toy and Harrison (1930) find such a drop in crystals which have been annealed for a certain time just below their melting-point, on the other hand, Lehfeldt (1935) finds that at − 185° C and in a field greater than 200 V/mm , one electron is drawn to the anode for every quantum absorbed

---

* The field produced by an electron at a distance of 10^{-4} cm is 10 V/cm

If the latter explanation is true we should expect that, for very weak intensities of illumination, the drop in efficiency would be less marked In fact, at low temperatures, above a certain intensity of illumination, we should not expect the rate of formation of silver to depend on the intensity of illumination at all At very low intensities, on the other hand, the temperature should have no effect, as in fig 3

Actually, however, the drop in efficiency may be due to both causes, and experiments on the print-out effect at different temperatures and intensities would enable them to be separated, if the results turn out to be similar to those illustrated in fig 3

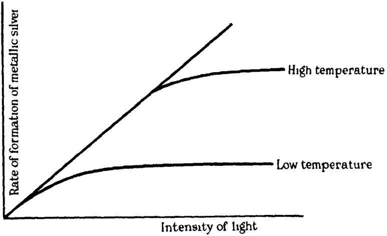


FIG 3—Rate of formation of silver in an emulsion (theoretical)

PART II—THE LATENT IMAGE

6—CONDITIONS FOR FORMATION

According to the investigations of Eggert and Noddack (1932) the number of quanta which a grain must absorb in order to become developable is of the order 100 This holds for the *average* grain of an emulsion, i e in an emulsion this number of quanta must be absorbed per grain in order to render about half the grains developable Of course in a sensitive emulsion there will be a small number of grains (about 1 %) which are sensitive to one or two quanta, there will also be other grains developable without exposure at all (causing fog)

Although many other theories have been put forward we shall assume that the latent image is a submicroscopic speck of metallic silver It is of course certain that the presence in the halide grain of metallic silver renders it developable, we know this from the simple fact that the developer continues to act on a grain long after any special configuration of atoms formed by the absorption of light has been obliterated

In discussing the latent image, then, we have to consider under what conditions a speck of metallic silver will start to grow on the surface of a halide grain In this connexion the "sensitivity specks" of Sheppard are of great interest According to him these are specks of silver sulphide adhering to the surface of the grain, and their function is to "concentrate" the silver atoms formed at definite points on the surface, where the developer can get at them From our point of view, what we require is a place on the surface where the *electrons* will "stick" If the conduction levels in silver sulphide are lower than in silver bromide, the sulphide speck will have just this effect, an electron in the conduction level of the latter, on meeting the sulphide speck will fall down a potential hill and become stuck (fig 4) The speck thus becomes negatively charged, silver ions are attracted to it, and a silver speck grows, as explained above

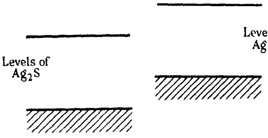


FIG 4—Conduction levels of AgBr and Ag_2S

We do not know the size of silver speck necessary for development, one might assume that since in the average grain 100 quanta must be absorbed, the speck must have 100 atoms It is, however, by no means certain *a priori* that the first few electrons released by the light are effective in producing silver atoms, and if we assume that the necessary size of silver speck is a good deal less than 100 atoms, and that the hundred quanta are necessary for the reason given below, a larger number of facts can be correlated, as we shall now see

Consider a speck of uncharged metallic silver in contact with AgBr Let E be the work necessary to remove an electron from the metal into the conduction level of the AgBr (cf fig 1) We have no *a priori* knowledge of E, but the considerations of § 7 show that it is much less than the work function of silver. Now as the temperature is raised electrons will pass from the metal into the conduction level of the salt. If the electron were uncharged we should get a "vapour" of electrons, the vapour pressure being proportional to $e^{-E/kT}$. If we denote by c the ratio of the number of electrons per unit volume in the conduction band to the number of ion pairs, then we should have, as regards the order of magnitude at any rate,

$$c \sim e^{-E/kT}$$

Actually the positive charge left on the metal will cause the negative electron gas to cluster round the surface of the metal, the concentration being of the order $e^{-E/kT}$ and dying away outside it.*

Now the condition for a silver speck to grow is that it shall be negatively charged, whereas we have just seen that a speck initially neutral gives off electrons and becomes positively charged. Clearly then, when a grain is exposed to light, a silver speck cannot begin to grow until the concentration c of electrons reaches the critical value c_0 given by

$$c_0 = e^{-E/kT}$$

Assuming that 100 quanta have to be absorbed by a grain of 10^{10} ion pairs in order that a silver speck shall start to grow, this gives $c_0 = 10^{-8}$ and

$$E = k \times 270^\circ \times 8 \log_e 10 \simeq 0.4 \text{ e-volt}$$

This is much less than the work function W, and suggests that the lowest conduction level in AgBr has negative energy, as in fig. 1.

The idea of a vapour pressure can hardly be applied to the silver speck in its initial stages of growth, when it would be unlikely to carry a charge (of either sign) numerically greater than e. In the presence of an electron vapour, however, such a speck would continuously be capturing and again losing an electron. If the grain is to grow the time average of the charge on it must be negative, and to achieve this at any finite temperature it is clear that a certain concentration of electrons must be reached in the halide speck.

7—THE RECIPROCITY LAW

As is well known the developed density D of a photographic image is not proportional to the product of the intensity I and the exposure time t, for given It, D in general falls off for low I The reason for this can be understood on the basis of the model given above

---

* The problem of an electron gas in thermal equilibrium with an electrode has been investigated in some detail (cf. for example, Fowler, 1936). The distance in which the density dies away is of the order $(kT/2\pi n_0 e^2)^{\frac{1}{2}}$, where n_0 is the number of electrons per cm.^{3} at the boundary. With $n_0 \sim 10^{15}$ this gives $3 \cdot 10^{-5}$ cm.

We have seen that the electron gas in the silver halide grain must reach a certain concentration $c_0 = e^{-E/kT}$ before the latent image can begin to form But while this concentration is forming the electrons can recombine with their original halogen atoms Eventually an equilibrium state is reached in which the electrons recombine as fast as they are formed If in this state the concentration is less than the critical concentration c_0, the latent image cannot begin to form

These ideas may be expressed analytically as follows Consider a single electron in a halide grain containing say 10^{12} ion pairs If it is on one of the six metal ions adjacent to a given halogen atom it will recombine with it after a time of the order of 10^{-8} sec On the other hand, during its motion through the crystal, it will on the average spend only a fraction 6×10^{-12} of its life in these six ions If then A is the probability per second that the electron and halogen atom recombine

$$A \sim 6 \times 10^{-4} \text{sec}^{-1}$$

Consider now a grain absorbing I_Q quanta per second, then after a time t the number N of electrons and of neutral halogen atoms is given by

$$\frac{dN}{dt} = I_Q - AN^2, \tag{1}$$

the rate of recombination being proportional to the product of the numbers of electrons and vacant places The solution of (1) is

$$N = \frac{I_Q}{A} \tanh\left[(I_Q A)^{\frac{1}{2}} t\right] \tag{2}$$

As $t \to \infty$, $N \to \sqrt{(I_Q/A)}$ Thus, however long the exposure, we cannot obtain more than a certain concentration of electrons In an ordinary exposure, a grain will absorb say 100 quanta/sec , and a maximum number of electrons of $2\ 5 \times 10^3$ can be reached with the above value of A For very weak intensities the maximum will be much less After the exposure the number n of electrons will decrease according to the law (obtained from (2) with $I_Q = 0$)

$$n = \frac{1}{At + 1/n_0}.$$

Fig 5 shows n during and after two exposures of equal It, but different I

In any emulsion the rate of absorption of light will vary from grain to grain, depending on the size and other factors But since for any grain a certain critical concentration must be reached before the latent image can form, we are driven to the conclusion that for any given intensity of illumination only a certain proportion of the grains will ever become developable, however long the exposure is continued This is in agreement with experiments by Toy (1926), who examined the grains of an emulsion microscopically after development Another way of expressing the same conclusion is to say that, to obtain a given density for given development, the intensity of illumination (I) must exceed a given value Fig 6 shows the type of result we should expect, we plot the value of It necessary to obtain

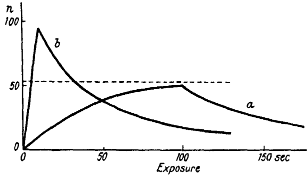


FIG 5—Calculated number (n) of electrons in a grain ($A = 4\ 10^{-4}$ sec $^{-1}$) (a) exposure of 100 sec, 1 quantum absorbed per sec (b) exposure of 10 sec, 10 quanta absorbed per sec

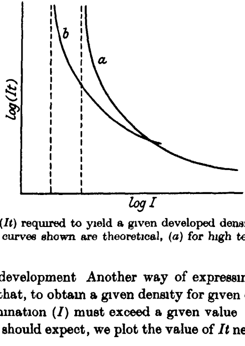


FIG 6—Exposure (It) required to yield a given developed density as a function of intensity (I) The curves shown are theoretical, (a) for high temperatures, (b) for low temperatures

a given density D on a developed plate, for varying I It will be seen that, for I below a given value, the density required cannot be obtained * According to the theory outlined above this minimum intensity should decrease rapidly with decreasing temperature, as shown by curve *b* in fig 6 Experiments to test this particular point do not seem to have been made, but its correctness is strongly suggested by some results of Webb (1935) on the breakdown of the reciprocity law at low temperatures, Webb's results are shown in fig 7 Webb does in fact find that for low intensities emulsions become *more* sensitive at low temperatures

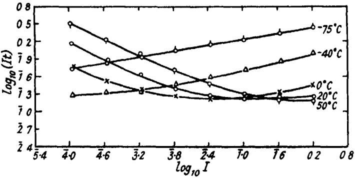


FIG 7—Exposure required to yield a given density (observed)

For high intensities it will be seen that raising the temperature increases the sensitivity For these intensities there is probably no difficulty in reaching the critical concentration of electrons, what is important is that the concentration should remain above the critical value long enough for electrolysis to take place and for the latent image to form The electrolytic conductivity is found to increase rapidly with increasing temperature (as $\exp(10{,}000/T)$), and this for high intensities seems to be the determining factor

Possibly the decrease in the sensitivity (fig 7) which occurs at low temperatures as the intensity is raised may be ascribed to the same cause At these temperatures and intensities there is no difficulty about reaching the critical concentration, the important thing is to give the electrolysis time to take place

Finally, some rather speculative remarks may be made about the few ultra-sensitive grains in fast emulsions which are sensitive to one or two quanta

---

* These conclusions depend on the assumption that the halogen atoms do not escape from the halide grain before recombination can take place

It is generally believed that the ripening process in the preparation of photographic emulsions leads to the formation of metallic silver in the grains Obviously this silver cannot in general be on the surface of the grains, or they would be developable (as a few are in over-ripened emulsions) If, however, a grain contains already a speck of silver in the interior, we have to enquire why this does not collect the negative charge released by the light and so grow in size, instead of a new silver speck being formed at the surface A possible answer seems to be that, if the mechanism sketched in fig 2 for the electrolytic conductivity is correct, and the current is carried by excess metal ions, the silver speck cannot grow because there is no room for it to do so A silver speck can only grow at the surface where the pressure caused by the adsorption of new silver ions can be relieved by the whole speck being pushed outwards

If these ideas are correct the effect of silver in the grain will be to increase the sensitivity, because the silver will give off electrons, and especially just near it there will already be a certain electron vapour pressure before exposure An ultra-sensitive grain will be one in which a sensitivity speck is situated near to a bit of silver embedded in the halide The superior speed of a coarse-grained emulsion will be due to the fact that in a large grain there is a greater surface area and hence a greater probability of such a state of affairs occurring

8—DEVELOPMENT

It is not our purpose to review the various theories of development We remark here only that if the negative ions of the developer can discharge electrons on to any metallic silver present on the surface, then electrolytic conductivity must take place in the solid crystal, in *exactly* the same way as when the electron is placed there by the action of light Development and photolysis may thus perhaps be regarded as giving rise to an exactly similar movement of ions within the halide crystal

In order, then, that a developer shall not reduce unexposed grains, we must assume the energy level of the negative ion to lie between the highest occupied level of the metallic silver and the lowest conduction level of the AgBr, or, if there are sensitivity specks, of Ag_2S

The validity of this hypothesis as to the nature of (chemical) development is independent of whether the developer is adsorbed to the surface of the latent image silver, as postulated by Rabinowitsch (1934)

9—HERSCHEL EFFECT

The developed density of an exposed plate can be lessened (though not brought back to the unexposed value) by exposing it to red light (cf Hay 1932*b*), the number of incident quanta of red light which are necessary is about 10^{10} times greater than the number of quanta of blue light necessary to produce the latent image

Let us assume that the primary process in the Herschel effect is the ejection of an electron from the latent image silver into the conduction levels of the silver-halide grain As we have seen, the energy $h\nu$ required to do this is quite low, well below that of red light Moreover, there are in an exposed grain about 10^{10} as many silver-halide molecules as atoms of silver in the latent image, and so, if the absorption coefficients of the two are of the same order of magnitude, we should expect that probabilities of absorption of a quantum by the latent image and by the halide grain to differ by a factor of this order

If the effect of the light is to remove an electron from the silver speck and so charge it positively, the latent image can only be destroyed because a positive ion is pushed after it by the resulting field The fact that the latent image can lose positive ions in this way suggests that it consists of very few atoms, one can hardly imagine positive ions being pushed out of a block of massive silver into silver halide

PART III—MECHANISM OF LIGHT ABSORPTION

10—DEPENDENCE ON TEMPERATURE

We have assumed up to the present that every absorbed quantum releases one electron into the conduction band, basing our assumption on the fact that photoconductivity is observed in large crystals According to experiments by Lehfeldt (1935) already referred to, each quantum absorbed by a silver bromide crystal illuminated at liquid air temperature releases one electron which can be drawn to the anode by an applied field

In the *alkali*-halides, on the other hand, this is not the case, no photoconductivity is observed for the characteristic absorption band (cf. for example Pohl 1937) The reason seems to be that in the first characteristic absorption band the light does not free an electron, but moves it only from a halogen ion to an unstable position in a neighbouring metal ion, where it is still in the field of the "positive hole", i e of the vacant place from which it came Thus if one illuminates in the tail of the characteristic absorption band, so that the light can penetrate well into the crystal, one does not free any electrons, if on the other hand one uses light of a shorter frequency so that electrons really are raised into the conduction band, the crystal is so opaque that only a layer 10^{-6} cm or so could be rendered photoconducting

With alkali-halide crystals most of the work on photoconductivity has been carried out by illuminating in the F-band, the F-band seems to be due to electrons trapped at a point where a negative ion is missing (cf Pohl 1937, discussion) Even in this case there is rather definite evidence that the light does not eject an electron *directly* into the conduction band, it raises it to an excited state, from which the thermal agitation of the crystal removes it into the condition band At low temperatures this does not happen before the electron has dropped back to its normal state, as we deduce from the fact (Pohl 1937) that the photo-current drops suddenly by a factor 100 as we lower the temperature below − 150° C

In silver halides one can hardly doubt that the absorption near its long-wave limit is due to the removal of an electron to an unstable position in a neighbouring ion only The photoconductivity observed can only be explained if we assume that thermal agitation then releases the electron into the conduction band If this is the case we should expect a drop in the photo-current at low temperatures, and for the same reason a drop in the sensitivity of emulsions

A large drop in the photo-current as the temperature is lowered to − 180° C has been observed by Toy and Harrison (1930) This result does not seem to be compatible with the more recent work of Lehfeldt (1935) mentioned above We are not then able to say at what temperature, if at all, the expected drop takes place

As Hertzfeld (1923) has pointed out, ions situated at the surface of cracks absorb light of lower frequency than ions in the body of the crystal Probably the long-wave end of the absorption spectrum of silver bromide is due to cracks, so that electrons will be released in a photographic emulsion mainly at the surface of the grains, for light of not too short a wave-length

11—DYE SENSITIZATION

The dye molecules are adsorbed to the surface of the grains Fig 8 shows the way the energy levels must be arranged to give sensitization in the red. the light is absorbed by the dye molecule raising an electron from the ground state A to an excited state B, if this lies above the conduction levels of AgBr, the electron can freely move from the dye molecule into the crystal, when latent image formation will take place as before

It has been shown that in the direct photochemical formation of silver in sensitized emulsions, one dye molecule can be active in forming as many as fifty silver atoms The dye molecule must thus be able to regain its electron In such emulsions halogen ions rather than halogen atoms will escape into the gelatine, and from these the dye molecule must regain its electron

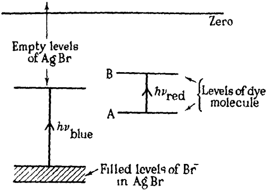


FIG 8

In dye-sensitized emulsions, as for ordinary emulsions, we should not expect the light to remove the electron directly into the conduction band, it will remove it into some excited state in which it is still in the field of the molecular ion Thermal motion will complete their separation Since the work necessary to separate them will probably be different from that to separate an electron and positive hole in the crystal, we may expect the temperature dependence of the process to be different from that in an unsensitized grain This is discussed further in the next section

12—LATENT IMAGE FORMATION AT LOW TEMPERATURES

A number of authors have investigated the dependence of developed density on the temperature at the time of exposure Sheppard, Wightman and Quirk (1934) find that the speed of a number of emulsions drops by about ten on lowering the temperature to that of liquid air They find no marked difference in the behaviours of ordinary and dye-sensitized emulsions Berg and Mendelssohn (1937) have extended these results to liquid hydrogen temperature Defining sensitivity as the exposure required to give a density of 0 1 above fog it was found for a blue sensitive emulsion the sensitivity at 90° K is 7 %, at 20° K 4 % of that at room temperature For a panchromatic emulsion exposed to red light the fall in the sensitivity was much greater, viz to 0 2 %

It is extremely improbable that any motion of ions can take place at 20° K, and the fact that any sensitivity at all survives at these temperatures suggests that the formation of the latent image silver speck takes place *after* the emulsion is warmed up If this is so, the electrons released by the light must become "stuck" in some way in the halide grain and released when it is warmed, a process which we know to occur at low temperatures in large crystals of silver bromide (cf § 3)

The drop in sensitivity may be due to two causes

1—There are not enough places where an electron can be stuck, so that some of the electrons released wander round in the crystal till they find their original halogen atoms, and recombine with them

2—The number of electrons released per absorbed quantum falls below unity as the temperature is lowered As we have seen, this is likely on theoretical grounds, and may well account for the difference between ordinary and dye-sensitized emulsions The rather contradictory evidence for unsensitized silver bromide has already been discussed above

In conclusion the authors would like to express their thanks to the staff of the research laboratory of Messrs Kodaks, Wealdstone, and also to Dr J Brentano for many discussions about the experimental material

SUMMARY

An attempt is made to explain the photolysis of silver halides in terms of the concepts of atomic physics The mechanism by which the silver atoms formed by the light coagulate to form specks of metallic silver is discussed The ideas used in this discussion are then turned to the photographic latent image, and are shown to account qualitatively for the variation of developed density with temperature, and, for given exposure, with intensity of light A brief discussion is given of the Herschel effect, and of sensitization by dyes

REFERENCES

Berg and Mendelssohn 1937 *Proc Phys Soc* **49** (extra number), 38
Eggert and Noddack 1923 *Z Phys* **20**, 299
— — 1932 *Handb wiss Photographie*, **5**, 263
Fowler 1936 "Statistical Mechanics", p 364

Frohlich 1937 *Proc Roy Soc* A, **160**, 230
Gläser and Lehfeldt 1936 *Nachr Ges Wiss Göttingen*, Math physikal Kl **1**, 91
Gurney and Mott 1937 *Proc Phys Soc* **49** (extra number), No 274, p 32
Hay 1932*a* *Handb wiss Photographie*, **5**, 65
— 1932*b* *Handb wiss Photographie*, **5**, 317
Hertzfeld 1923 *Z phys Chem* **105**, 329
Hilsch and Pohl 1930 *Z Phys* **64**, 606
Jost 1937 "Diffusion und chemische Reaktion in festern Stoffen" Dresden
Lehfeldt 1933 *Z Phys* **85**, 717
— 1935 *Nachr Ges Wiss Göttingen*, Math physikal Kl **1** 170
Lohle 1933 *Nachr Ges Wiss Göttingen*, p 271
Mott 1937 *Proc Phys Soc* (in Press)
Pohl 1937 *Proc Phys Soc* **49** (extra number), 1
Rabinowitsch 1934 *Z wiss Phot* **33**, 57
Sheppard, Wightman and Quirk 1934 *J Phys Chem* **38**, 817
Toy 1926 *Brit J Photogr* p 704
Toy and Harrison 1930 *Proc Roy Soc* A, **127**, 613
Webb 1935 *J Amer Opt Soc* **25**, 4
— 1936 *J Amer Opt Soc* **26**, 367

Self-consistent Field with Exchange for Calcium

By D R Hartree, F R S and W Hartree

(*Received* 22 *October* 1937)

1—Introduction

The solution of Fock's equation (Fock 1930) of the self-consistent field with exchange has now been carried out for several atoms (Fock and Petrashen 1934, 1935, D R and W Hartree 1935 *b*, 1936 *a*, *b*, *c*) The present paper gives the results of similar calculations for Ca^{++}, and their extension to some of the more important states of Ca^+ and neutral Ca

Calcium was chosen as the next element for which to do such calculations for two reasons Firstly, some of the lines in the arc and spark spectra are important in astrophysical applications, for which transition probabilities are required, it seemed likely that the calculated transition probabilities might be appreciably affected by effects of exchange terms on the wave functions from which they were evaluated, and on account of their astrophysical applications it seemed desirable to obtain the best approximations to these transition probabilities which are practicable at present

Secondly, it seemed likely that from the differences between the results for the self-consistent field with and without exchange, for Cl^- and Ca^{++}, it would be possible to interpolate the corresponding differences for K^+ and Ar with adequate accuracy for most purposes Results for K^+ were required for calculating X-ray scattering factors more accurate than those derived from the results of the self-consistent field without exchange Also it seemed that the calculation of the diamagnetic susceptibility of argon, and of the difference between the susceptibilities of Na^+ and K^+, from the solutions of Fock's equation for the respective ions, and the comparison of these results with the observed values, might throw some light on the differences between the calculated and observed susceptibilities of ions Further the solution of Fock's equations for argon may at any time be required for calculation of electron scattering, or of the Auger effect

It will appear later that this interpolation for K^+ and Ar cannot be made as certainly as was anticipated, but still it seemed that it should give a fair approximation to the effect of exchange, and should provide fairly good sets of estimates for starting the solutions of Fock's equations for K^+ and Ar

2—FOCK'S EQUATIONS FOR ELECTRONS OUTSIDE CLOSED GROUPS

A convenient derivation of Fock's equations for a configuration consisting of closed (nl) groups, by formal differentiation of the expression for energy integral $E = \int \Psi^* H \Psi d\tau \Big/ \int \Psi^* \Psi d\tau$, in terms of radial wave functions, with respect to these wave functions has already been given (D R and W Hartree 1936 *b*), and we need only consider the additional features arising from the electrons outside closed groups

If E is expressed in terms of the integrals I, F, G (in the notation of the authors' previous papers, see particularly 1936 *b*), the contributions to E can be divided into three groups, namely

a—Those contributions arising from the core of closed groups only

b—Those arising from the I integrals for the wave functions outside the core, and the F and G integrals for the interactions between the electrons in each of these wave functions and those of the core

c—The F and G integrals arising from the interactions between the electrons in the outer wave functions

We will consider first a single electron outside a core of closed groups, and write (nl) for the wave function of the outer electron, and $(n'l')$ for those of the core

The contributions to the energy in group a do not give any terms in Fock's equation for the (nl) wave function, and, for one outer electron only, there are no contributions in group c The contributions in group b can be obtained as follows It is known that both the Coulomb and exchange interactions between an electron in an (nl) wave function and a complete $(n'l')$ group are independent of the other quantum numbers of the (nl) wave function (Condon and Shortley 1935, Chap 6, § 9, formula (11)) and hence are $1/2(2l+1)$ times those for the interaction between complete (nl) and $(n'l')$ groups These contributions to the interaction between complete groups have been tabulated in a previous paper (D R and W Hartree 1935 *b*, Table I *b*), and using these results, we have for the contributions in group b, for a single electron outside closed groups,

$$I(nl)+\Sigma_{n'l'}\left[2(2l'+1)\,F_0(nl,n'l')-\frac{1}{2(2l+1)}\Sigma_k B_{ll'k}G_k(nl,n'l')\right] \tag{1}$$

where the coefficients $B_{ll'k}$ are those previously tabulated

Now $I(nl)$ is also $1/2(2l+1)$ times the corresponding contribution to E from a complete (nl) group, hence the ratios of the coefficients of the terms in (1) are the same as for a complete (nl) group Thus, when Fock's equation for the (nl) radial wave function $P(nl)$ is formed by formal differentiation of (1) with respect to $P(nl)$, the coefficients of the terms arising from Coulomb and exchange interaction between the outer electron and the core are the same as for a complete (nl) group already given (D R and W Hartree 1936*b*, formulae (12)–(14)) There are now no terms arising from interaction within an (nl) group, hence Fock's equation for a single (nl) wave function outside closed groups is

$$\Sigma_{n'l'}\,H'_{nl\,n'l'}\,P(n'l'\mid r)=0, \tag{2}$$

where

$$H'_{nl\,nl}=\frac{d^2}{dr^2}+\frac{2}{r}\,[N-\Sigma_{n'l'}\,2(2l'+1)\,Y_0(n'l',n'l'\mid r)]-\epsilon_{nl\,nl}-\frac{l(l+1)}{r^2} \tag{3}$$

$$=\frac{d^2}{dr^2}+\frac{1}{r}\,T(r)-c_{nl\,nl}-\frac{l(l+1)}{r^2}, \tag{4}$$

$T(r)$ being the total $2Z_p$ of the field of the nucleus and the Schrodinger charge distribution of the core, namely (cf D R and W Hartree 1936*b*, p 51)

$$\begin{aligned} T(r)&=2[N-\Sigma_{n'l'}\,2(2l'+1)\,Y_0(n'l',n'l'\mid r)]\\ &=2C+\Sigma_{n'l'}\,4(2l'+1)\,[1-Y_0(n'l',n'l'\mid r)], \end{aligned} \tag{5}$$

where C is the net charge on the core, and

$$H'_{nl\;nl} = \Sigma_k \frac{B_{llk}}{2l+1}\frac{Y_k(nl, n'l)}{r} - \epsilon_{nl,\,nl} \quad (n' \neq n), \tag{6}$$

$$H'_{nl\;nl} = \Sigma_k \frac{B_{ll\,k}}{2l+1}\frac{Y_k(nl, n'l')}{r} \quad (l' \neq l), \tag{7}$$

the coefficients $B_{ll\,k}$ being those already tabulated (D R and W Hartree 1936*b*, Table I *b*)

For a confirmation of two or more electrons outside closed groups, the contributions to the energy in group *b* similarly give the same contributions (3), (6), (7) to Fock's equations for the outer wave functions In addition there are the contributions to the energy in group *c* and the corresponding terms in Fock's equations For many simple configurations these contributions to the energy can be written down at once from the results given by Slater (1929) or their extension by Condon and Shortley (1935 Chap 7, § 5), or, if not already available, they can be obtained by Slater's method (1929) for any configuration which gives only one term for each (L,S) From the contributions to the energy the contributions to Fock's equations can be written down by formal differentiation with respect to each (*nl*) radial wave function using results already given (D R and W Hartree 1936*b*, formulae (7)–(10)), and dividing by the number of corresponding (*nl*) wave functions (if more than one)

For approximate results, it will probably be sufficiently accurate to neglect the perturbation of the core by the outer electrons If it is required to take this perturbation into account, the contribution to Fock's equations for the core wave functions, arising from interaction between the core and the outer electrons, are found by differentiating the energy contributions in group *b* with respect to the core wave functions It is found that in Fock's equation for the $(n'l')$ core wave function, each (*nl*) outer wave function contributes terms which are $1/2(2l+1)$ times those for a complete group

For a configuration of complete groups, it is always possible to make linear transformations of the radial wave functions so as to reduce the non-diagonal Lagrange multipliers $\epsilon_{nl\;nl}(n' \neq n)$ in (4) to zero But, as pointed out for a simple case in a previous paper (D R and W Hartree 1936*a*, pp 594–5), for a configuration including some incomplete groups this is not in general possible, since if either the (*nl*) or the $(n'l)$ group is incomplete, a linear transformation of the radial wave functions does not leave unchanged the determinant for the wave function of the whole atom

Hence only those non-diagonal parameters $\epsilon_{nl\,n''l}$ for which $(n'l)$ and $(n''l)$ are both core wave functions can be reduced to zero, those involving an outer wave function must be retained and adjusted to satisfy the conditions of orthogonality of the outer wave functions to the core wave functions and to one another

3—PROCESS OF SOLUTION OF FOCK'S EQUATIONS FOR CALCIUM

Solutions of Fock's equations have been carried out for the normal state of the argon-like ion Ca^{++}, for the $(4s)$, $(4p)$, and $(3d)$ states of Ca^+, and for the $(4s)^2\,{}^1S$ and $(4s)(4p)\,{}^3P$ and ${}^1P$ states of neutral Ca The three states of Ca^+ for which calculations were made are the three lowest states, and the most important for the astrophysical application of the results, and also the most interesting physically from the point of view of the relative energies of $(4s)$ and $(3d)$, and the relation of this to the place in the periodic table at which the transition elements of the iron group begin For the $(4s)(4p)$ configuration of the neutral atom, the evaluation of the solution of the equations is easier for the ${}^3P$ state, but the ${}^1S$–${}^1P$ transition gives an astrophysically important line ($\lambda = 4227$), and it seemed worth while to attempt the calculation of the ${}^1P$ state also

The procedure used for the evaluation of the solution of Fock's equations for Ca^{++} was very similar to that already used for Cl^- and Cu^+ (D R and W Hartree 1936*b*, § 3, and 1936*c*, § 2), the equations are the same as for Cl^- except for the change in the value of the atomic number N The ionic charge $+2$, and the consequent comparative insensitiveness of the wave functions to small changes in the estimates, made the work markedly easier than that for the more sensitive structures Cl^- and Cu^+ The initial estimates of the $Z_k(nl, n'l')$ functions were based on the result of the self-consistent field without exchange (D R and W Hartree 1935*a*) differences between these Z functions for the self-consistent field with and without exchange were estimated roughly from the corresponding differences for Cl^- and Cu^+ Four successive approximations (one concerned with $(3p)$ only) were required to obtain results self-consistent to the order of accuracy attained in the solution of Fock's equation for other atoms For Ca^+ and neutral Ca, the perturbation of the Ca^{++} core by the outer electrons was neglected, as the amount of work involved in the calculation of the outer wave functions was already very considerable this perturbation is likely to be small, except perhaps for configurations involving $(3d)$

For the self-consistent field without exchange, the calculation of the (nl) wave function of a single series electron in the field of the core regarded as given needs no estimates (except that of $\epsilon_{nl\ nl}$) But since in Fock's equations the $Z_k(nl, n'l')$ functions in the exchange terms between the (nl) wave function and the core involve this wave function, it (or the set of exchange terms involving it) has to be estimated before the integration is begun *

For each state of Ca^+ the outer (nl) wave function, calculated without exchange, was taken to provide the first estimates of the various $Z_k(nl, n'l')$ functions, and these were subsequently improved by a process of successive approximation

In the derivation of Fock's equation for an (nl) outer wave function, it is supposed that this wave function is orthogonal to the core $(n'l)$ wave function with the same l, and the non-diagonal Lagrange multipliers $\epsilon_{nl\ n'l}$ are introduced to enable this condition to be satisfied Actually it was found unnecessary to use non-zero values of these parameters, either for $(4s)$ or $(4p)$, since the wave functions calculated without them were already orthogonal to the relevant core wave functions, to the accuracy of the calculations For $(3d)$ there are no core wave functions with the same l, so the question of orthogonality to them does not arise

For neutral Ca, the $(4s)^2$ configuration consists of closed groups and there are no exchange interactions within the $(4s)^2$ group, hence the equation can be derived from that for Ca^+ $(4s)$ by the addition of the term in $T(r)$ (cf formula (5)), arising from the Coulomb interaction between the two $(4s)$ electrons

For the configuration $(4s)$ $(4p)$, the terms in the equation arising from the interaction between the $(4s)$ and $(4p)$ electrons can be written down from the corresponding terms for the $(2s)$ $(2p)$ configuration of Be, by suitable changes of the quantum numbers Thus, for normalized wave functions, these equations are

$$\left[\frac{d^2}{dr^2}+\frac{T(\mathrm{Ca}^{++})-2Y_0(4p,4p)}{r}-\epsilon_{4s\ 4s}\right]P(4s)$$
$$+\sum_{n=1}^{3}\left[\frac{2Y_0(ns,4s)}{r}-\epsilon_{ns\ 4s}\right]P(ns)+\sum_{n=2}^{3}\frac{2Y_1(np,4s)}{r}P(np)$$
$$\pm\frac{2}{3}\frac{Y_1(4s,4p)}{r}P(4p)=0, \qquad (8)$$

---

* Alternatively, Fock's equation for the (nl) outer wave function can be solved as an integro differential equation, this process would involve the evaluation of a definite integral over the range 0 to ∞ for each step of the integration, and seems more elaborate than solution by successive approximation, using estimates of the $Z_k(nl, n'l')$ functions Another alternative would be to use a modification of Torrance's method (Torrance 1934, for the application of this method to Fock's equation, see D R and W Hartree 1935*b*, § 6)

$$\left[\frac{d^2}{dr^2}+\frac{T(\mathrm{Ca}^{++})-2Y_0(4s,4s)}{r}-\epsilon_{4p,4p}-\frac{2}{r^2}\right]P(4p)$$
$$+\frac{2}{3}\sum_{n=1}^{3}\frac{Y_1(ns,4p)}{r}P(ns)+\sum_{n=2}^{3}\left[\frac{2Y_0(np,4p)}{r}+\frac{4}{5}\frac{Y_2(np,4p)}{r}-\epsilon_{np,4p}\right]P(np)$$
$$\pm\frac{2}{3}\frac{Y_1(4s,4p)}{r}P(4s)=0, \qquad (9)$$

the upper sign referring to the triplet and the lower to the singlet state

The process of solution of these equations followed much the same lines as that for the corresponding states of Be, apart from the use of normalized wave functions throughout the present work On account of the much larger number of $Z_k(nl, n'l')$ functions expressing the interaction of each (nl) outer wave functions with the $(n'l')$ wave function of the core, the work was much heavier than for Be, even though the work for Be gave some indication of the likely effects of exchange terms, and helped in making the initial estimates As for Be, the calculations were much more troublesome for the ^{1}P state than for the ^{3}P state, on account of the repulsive effect of the exchange terms for the singlet state The numbers of steps of successive approximation were three, five, and eight for the $(4s)^2$, $(4s)\,(4p)\,^3$P and $(4s)\,(4p)\,^1$P states respectively

It was not found possible to make the outer wave functions orthogonal to those of the core without introducing non-zero values of the non-diagonal Lagrange multipliers, though the values required are small Their introduction makes the work more troublesome, and the course of the work showed that there was no object in introducing them until a close approach to a self-consistent solution has been obtained without them

4—RESULTS OF Ca^{++}

The results for Ca^{++} are given in Tables I–III in the same form as for other atoms for which similar calculations have been made Table I gives the normalized radial wave functions for Ca^{++}, the third decimal in the tabulated values of $P(nl)$ should be reliable to a unit throughout Table II gives the corresponding contributions to Z and the total $2Z_p$, and Table III the radial charge density $U(r)$ calculated with and without exchange, and the difference

Comparison of the differences between the wave function calculated with and without exchange, with the corresponding differences for Cl^- showed one surprising feature

TABLE I—Ca^{++} NORMALIZED WAVE FUNCTIONS

Table of P_N/r^{l+1}

r	(1s)	(2s)	(2p)	(3s)	(3p)
0	174 7$_5$	50 73	248 3	17 45$_5$	81 34
0 005	158 1	45 86	236 2	15 78	77 40
0 010	143 1	41 36	224 8	14 22$_5$	73 66
0 015	129 5$_5$	37 21	214 0	12 79	70 10
0 020	117 3	33 39	203 8	11 46$_5$	66 72
0 025	106 2	29 86	194 1	10 24	63 52
0 030	96 2	26 60	184 9	9 11	60 48

Table of P_N

r	(1s)	(2s)	(2p)	(3s)	(3p)
0	0 000	0 000	0 000	0 000	0 000
0 01	1 431	0 414	0 022$_5$	0 142	0 007
0 02	2 346	0 668	0 081	0 229	0 027
0 03	2 885	0 798	0 166	0 273	0 054
0 04	3 156	0 834	0 269	0 284	0 088
0 05	3 240	0 799	0 382	0 270	0 125
0 06	3 194	0 711	0 501	0 238	0 163
0 07	3 063	0 586	0 622	0 192	0 201
0 08	2 880	0 435	0 741	0 137	0 239
0 10	2 439	0 094	0 967	0 014	0 308
0 12	1 986	−0 260	1 166	−0 111	0 366
0 14	1 576	−0 593	1 333	−0 227	0 411
0 16	1 227	−0 890	1 466	−0 326	0 442
0 18	0 943	−1 141	1 566	−0 404	0 458
0 20	0 716	−1 344	1 636	−0 463	0 462
0 22	0 540	−1 500	1 678	−0 500	0 454
0 24	0 404	−1 614	1 696	−0 517	0 436
0 26	0 301	−1 688	1 694	−0 517	0 408
0 28	0 224	−1 730	1 675	−0 501	0 372
0 30	0 166	−1 743	1 641	−0 471	0 330
0 35	0 078	−1 683	1 515	−0 350	0 204
0 40	0 037	−1 538	1 352	−0 188	0 061
0 45	0 017	−1 354	1 177	−0 006	−0 089
0 50	0 008	−1 159	1 005	+0 178	−0 235
0 55	0 003$_5$	−0 972	0 846	0 353	−0 371
0 60	0 001$_5$	−0 801	0 704	0 511	−0 495
0 7	—	−0 526	0 473	0 763	−0 697
0 8	—	−0 334	0 310	0 927	−0 836
0 9	—	−0 207	0 200	1 012	−0 918
1 0	—	−0 127	0 127	1 036	−0 954
1 1	—	−0 077	0 080	1 015	−0 955
1 2	—	−0 047	0 050	0 963	−0 929
1 4	—	−0 017$_5$	0 020	0 814	−0 832
1 6	—	−0 007	0 008	0 651	−0 708
1 8	—	−0 003	0 003	0 501	−0 582
2 0	—	−0 001	0 001$_5$	0 375	−0 466

TABLE I—*continued*

Table of P_v

r	$(1s)$	$(2s)$	$(2p)$	$(3s)$	$(3p)$
2 2	—	$-0\ 000_5$	$0\ 000_5$	0 275	$-0\ 367$
2 4	—	—	—	0 199	$-0\ 285$
2 6	—		—	0 142	$-0\ 219$
2 8	—	—	—	0 100	$-0\ 166$
3 0	—	—	—	0 070	$-0\ 125$
3 2	—	—	—	0 049	$-0\ 094$
3 4	—			0 033	$-0\ 070$
3 6	—	—	—	0 023	$-0\ 052$
3 8	—	—	—	$0\ 015_5$	$-0\ 038$
4 0	—	—		$0\ 010_5$	$-0\ 028$
4 5	—	—	—	0 004	$-0\ 013$
5 0	—			$0\ 001_5$	$-0\ 006$
5 5	—	—	—	$0\ 000_5$	$-0\ 002_5$
6 0	—	—	—	—	$-0\ 001$
6 5	—	—	—	—	$-0\ 000_5$
$\epsilon_{nl\ nl}$	299 9	34 75	28 37	5 557	3 756

Considering a given configuration for atoms of different atomic number N, and comparing corresponding radii in the different atoms (for example, the radii at the principal maximum of $|P(3p)|$), it would be expected that the effective nuclear charge Z at this point would vary like $N-s$, where s is some screening number nearly independent of N, whereas the non-diagonal "exchange" Z_k functions would be approximately the same, hence the relative importance of the exchange and Coulomb forces would be expected to vary as $1/(N-s)$, so that the effect of exchange terms would increase with decreasing N and so be greater for Cl^- than for Ca^{++} This expectation is fulfilled for all the core wave functions except for the $(3s)$ wave function in the neighbourhood of its main maximum, over a considerable range in this region the difference between the wave functions, calculated with and without exchange, for Cl^- is numerically smaller than for Ca^{++} or even of the opposite sign This result seemed at first to suggest an error in the results for Cl^-, but later work on K^+ and Ar, of which the results will be given in a separate paper, confirmed this unexpected behaviour

This behaviour is another aspect of the small net effect of exchange on the $(3s)$ wave function of Cl^-, of which a qualitative explanation has already been given (D R and W Hartree 1936*b*, p 58) It suggests that the variation, with atomic number N, of the differences between the wave functions calculated with and without exchange is far from linear in this region, and a considerable departure from linearity may also be expected

TABLE II—Ca^{++} CONTRIBUTIONS TO Z AND TOTAL $2Z_p$

r	$2(2l+1)\,[1-Z_0(nl, nl \mid r)]$					Total
	(1*s*)	(2*s*)	(2*p*)	(3*s*)	(3*p*)	$2Z_p$
0	2 000	2 000	6 000	2 000	6 000	40 00
0 005	1 998	2 000	6 000	2 000	6 000	$39\ 20_5$
0 010	1 985	1 999	6 000	2 000	6 000	38 43
0 015	1 956	1 996	6 000	2 000	6 000	$37\ 67_5$
0 020	1 910	1 992	6 000	1 999	6 000	$36\ 94_5$
0 025	1 847	1 988	6 000	1 999	6 000	$36\ 24_5$
0 030	1 770	1 982	5 999	1 998	6 000	35 57
0 04	1 585	1 968	5 996	1 996	6 000	34 32
0 05	1 379	1 954	5 990	1 995	5 999	33 18
0 06	1 170	1 943	5 978	1 993	5 998	32 13
0 07	0 973	1 934	5 959	1 992	5 996	31 16
0 08	0 797	1 929	5 931	1 992	5 993	$30\ 25_5$
0 10	0 512	1 926	5 843	1 992	5 984	28 59
0 12	0 315	1 925	5 706	1 991	5 970	$27\ 07_5$
0 14	0 188	1 917	5 517	1 990	5 952	$25\ 67_5$
0 16	0 110	1 895	5 281	1 987	5 930	24 37
0 18	0 063	1 853	5 004	1 982	5 905	$23\ 16_5$
0 20	0 035	1 791	4 696	1 974	5 880	$22\ 04_5$
0 22	0 020	1 709	4 365	1 965	5 855	21 02
0 24	0 011	1 612	4 023	1 955	5 831	$20\ 07_5$
0 26	0 006	1 503	3 678	1 944	5 809	19 21
0 28	0 003	1 385	3 336	1 933	5 791	$18\ 42_5$
0 30	$0\ 001_5$	1 265	3 006	1 924	5 776	17 71
0 35	$0\ 000_5$	0 968	2 254	1 907	5 754	$16\ 18_5$
0 40	—	0 707	1 635	1 899	5 748	$14\ 93_5$
0 45	—	0 497	1 155	1 898	5 747	13 89
0 50	—	0 338	0 797	1 897	5 739	$12\ 97_5$
0 55	—	0 225	0 540	1 889	5 711	$12\ 16_5$
0 60	—	0 147	0 360	1 870	5 654	11 41
0 7	—	0 059	0 154	1 786	5 436	$10\ 06_5$
0 8	—	0 023	0 064	1 641	5 079	8 89
0 9	—	0 008	0 026	1 451	4 613	7 88
1 0	—	0 003	0 010	1 239	4 084	$7\ 03_5$
1 1	—	0 001	0 004	1 028	3 534	$6\ 34_5$
1 2	—	—	$0\ 001_5$	0 831	3 000	5 79
1 4	—	—	—	0 511	2 059	$5\ 01_8$
1 6	—	—	—	0 296	1 343	$4\ 56_3$
1 8	—	—	—	0 163	0 843	$4\ 30_4$
2 0	—	—	—	0 087	0 513	$4\ 16_3$
2 2	—	—	—	0 045	0 305	$4\ 08_4$
2 4	—	—	—	0 023	0 177	$4\ 04_4$
2 6	—	—	—	0 011	0 102	$4\ 02_3$
2 8	—	—	—	$0\ 005_5$	0 057	$4\ 01_1$
3 0	—	—	—	$0\ 002_5$	0 032	$4\ 00_6$
3 2	—	—	—	0 001	$0\ 017_5$	$4\ 00_3$
3 4	—	—	—	$0\ 000_5$	0 009	$4\ 00_1$
3 6	—	—	—	—	0 005	$4\ 00_0$
3 8	—	—	—	—	$0\ 002_5$	$4\ 00_0$
4 0	—	—	—	—	0 001	$4\ 00_0$

TABLE III—Ca^{++} CHARGE DISTRIBUTION TOTAL RADIAL DENSITY $U(r) = \Sigma_{nl} 2(2l+1) P_N^2(nl \mid r)$ CALCULATED BY SELF-CONSISTENT FIELD (a) WITHOUT EXCHANGE, (b) WITH EXCHANGE, AND THE DIFFERENCE $(b)-(a)$

r	$U(r)$ (a)	$U(r)$ (b)	$\delta U(r)$ $(b)-(a)$
0·000	0	0	0·00
0·005	1·37	1·37	0·00
0·010	4·47	4·48	0·01
0·015	8·25	8·26	0·01
0·020	12·03	12·05	0·02
0·025	15·42	15·44	0·02
0·030	18·23	18·25	0·02
0·04	21·93	21·96	0·03
0·05	23·35	23·39	0·04
0·06	23·15	23·20	0·05
0·07	22·02	22·09	0·07
0·08	20·54	20·64	0·10
0·10	17·86	18·08	0·22
0·12	16·63	17·01	0·38
0·14	16·93	17·45	0·52
0·16	18·23	18·87	0·64
0·18	19·97	20·68	0·71
0·20	21·67	22·40	0·73
0·22	23·01	23·71	0·70
0·24	23·85	24·47	0·62
0·26	24·12	24·63	0·51
0·28	23·88	24·25	0·37
0·30	23·17	23·39	0·22
0·35	20·04	19·93	−0·11
0·40	16·09	15·79	−0·30
0·45	12·34	12·01	−0·33
0·50	9·36	9·14	−0·22
0·55	7·27	7·25	−0·02
0·60	6·04	6·24	+0·20
0·7	5·39	5·97	0·58
0·8	5·95	6·71	0·76
0·9	6·68	7·43	0·75
1·0	7·13	7·74	0·61
1·1	7·17	7·58	0·41
1·2	6·85	7·06	0·21
1·4	5·60	5·49	−0·11
1·6	4·16	3·85	−0·31
1·8	2·90	2·53	−0·37
2·0	1·94	1·59	−0·35
2·2	1·25$_5$	0·96	−0·29$_5$
2·4	0·79$_5$	0·57	−0·22$_5$
2·6	0·49$_5$	0·33	−0·16$_5$
2·8	0·30	0·18$_5$	−0·11$_5$
3·0	0·18	0·10$_5$	−0·07$_5$
3·2	0·11	0·06	−0·05
3·4	0·06$_5$	0·03	−0·03$_5$
3·6	0·04	0·02	−0·02
3·8	0·02	0·01	−0·01
4·0	0·01	0·00$_5$	−0·00$_5$
4·2	0·00$_5$	—	−0·00$_5$

in the outer part of the $(3p)$ wave function, though it is not shown in the same way, this non-linearity is the reason for the difficulty, mentioned in § 1, of interpolating these differences for K$^+$ and Ar from those for Ca^{++} and Cl$^-$

5—RESULTS FOR Ca$^+$ WAVE FUNCTIONS

Table IV gives the normalized radial wave functions $P(nl \mid r)$ for the three states of Ca$^+$, namely $(4s)$, $(4p)$ and $(3d)$, for which calculations have been made, the third decimal in the tabulated values should be reliable to 1 unit The wave functions, calculated with and without exchange, are shown

TABLE IV—Ca$^+$ NORMALIZED WAVE FUNCTIONS FOR SERIES ELECTRON

Table of P_N/r^{l+1}

r	$(4s)$	$(4p)$	$(3d)$
0	4 852	19 51	23 49
0 005	4 385	18 56	22 72
0 010	3 953	17 66	21 98
0 015	3 554	16 80$_5$	21 27
0 020	3 185$_5$	15 99$_5$	20 60
0 025	2 845	15 23	19 95
0 030	2 530	14 50	19 33

Table of P_N

0	0	0	0
0 01	0 039$_5$	0 002	0 0000
0 02	0 064	0 006$_5$	0 0002
0 03	0 076	0 013	0 0005
0 04	0 079	0 021	0 0012
0 05	0 075	0 030	0 0021
0 06	0 066	0 039	0 0035
0 07	0 053	0 048	0 0052
0 08	0 038	0 057	0 0074
0 10	0 003	0 073	0 013
0 12	−0 032	0 087	0 020
0 14	−0 064	0 098	0 028
0 16	−0 091	0 105	0 038
0 18	−0 113	0 108	0 049
0 20	−0 128	0 109	0 061
0 22	−0 138	0 107	0 074
0 24	−0 142	0 102	0 088
0 26	−0 141	0 094	0 102
0 28	−0 135	0 085	0 116
0 30	−0 126	0 075	0 131
0 35	−0 090	0 043	0 170
0 40	−0 042	0 008	0 208
0 45	+0 011	−0 029	0 246
0 50	0 063	−0 064	0 282
0 55	0 112	−0 097	0 317
0 60	0 154	−0 125	0 350
0 7	0 216	−0 168	0 407
0 8	0 246	−0 194	0 454
0 9	0 247	−0 202	0 492
1 0	0 226	−0 196	0 521
1 1	0 188	−0 179	0 542
1 2	0 139	−0 153	0 557
1 4	0 021	−0 085	0 574
1 6	−0 103	−0 004	0 577
1 8	−0 220	+0 079	0 572
2 0	−0 322	0 160	0 561
2 2	−0 408	0 235	0 546

TABLE IV—*continued*

Table of P_N

r	$(4s)$	$(4p)$	$(3d)$
2.4	−0.475	0.302	0.528
2.6	−0.525	0.361	0.508
2.8	−0.559	0.410	0.487
3.0	−0.579	0.451	0.464
3.2	−0.586	0.481	0.441
3.4	−0.584	0.504	0.417
3.6	−0.574	0.519	0.393
3.8	−0.557	0.527	0.369
4.0	−0.535	0.529	0.346
4.5	−0.466	0.511	0.289
5.0	−0.389	0.471	0.238
5.5	−0.315	0.419	0.193
6.0	−0.249	0.363	0.154
6.5	−0.193	0.307	0.121
7.0	−0.147	0.254	0.095
8	−0.081	0.168	0.056
9	−0.043	0.105	0.032
10	-0.022_5	0.064	0.018_5
11	−0.011	0.038	0.010
12	-0.005_5	0.021_5	0.005_5
14	-0.001_5	0.006_5	0.001_5
16	-0.000_5	0.002	0.000_5
18	—	0.000_5	—
$\epsilon_{nl,\,nl}$	0.8295	0.6193	0.6659

graphically in fig. 1. This figure shows strikingly the very large effect of the exchange terms on the $(3d)$ wave function, compared with their effect on the $(4s)$ and $(4p)$ functions, for $(4s)$ and $(4p)$ the exchange terms form a comparatively small proportion of $d^2P(nl)/dr^2$ over most of the range, whereas for $(3d)$ they form the main contribution to d^2P/dr^2 in the neighbourhood of the main maximum, so that the shapes of the curves of $P(3d)$, calculated with and without exchange, are quite different.

The inset in fig. 1 shows the charge distribution of the Ca^{++} ion on the same scale of r, and is included to show how the radial scale of the $(3d)$ wave function is related to that of the core, as calculated without exchange, the $(3d)$ wave function lies mainly well outside the core, corresponding to a "non-penetrating" orbit of the orbital atomic model, the effect of exchange is to pull it in considerably and make its dimensions very much more like those of the core wave functions with $n = 3$.

Another aspect of this large effect of exchange on the $(3d)$ wave function is shown by the following values referring to the normalized wave functions at the origin

	(4*s*)	(4*p*)	(3*d*)
$\mathrm{Lt}_{r\to 0}\left[\frac{P(\text{s c f with exchange})}{P(\text{s c f without exchange})}\right]$	0 932	0 833	2 182

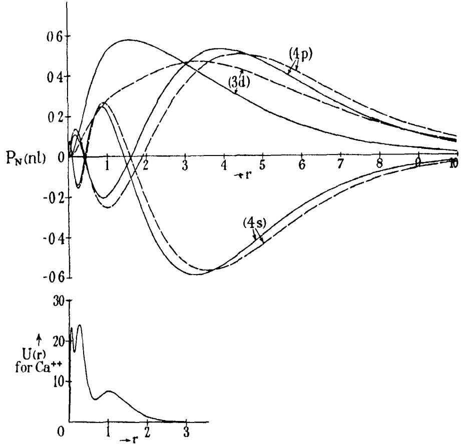


FIG 1—Normalized radial wave functions for series electron for (4*s*), (4*p*) and (3*d*) states of Ca⁺, calculated with exchange (full line) and without exchange (broken line) Inset below, radial distribution of charge of Ca⁺⁺ core, on same scale of *r*

Since spin and hyperfine structure separations depend mainly on the normalized P^2 for small r, values of these quantities for (3*d*), calculated from the wave functions with and without exchange, would differ by a factor of nearly 5

6—RESULTS FOR Ca⁺. ENERGY VALUES

Consider an atom consisting of a core of closed shells and a single series electron in an (nl) wave function. If E_{core} is the contribution to the energy from the I, F, G integrals involving core wave functions only, the total energy of the system is

$$E_{\text{core}} + I(nl) + \Sigma_{n'l'}\left[2(2l'+1)F_0(nl,n'l') - \frac{1}{2(2l+1)}\Sigma_k B_{ll'k}\, G_k(nl,n'l')\right],$$

whereas that of the ionized system consisting of the core only is just E_{core}. If the perturbation of the core by the series electron is neglected, the two values of E_{core} are exactly equal and the energy E_1 of the [core + (nl)] state, relative to the normal state of the core as zero, is

$$E_1 = I(nl) + \Sigma_{n'l'}\left[2(2l'+1)F_0(nl,n'l') - \frac{1}{2(2l+1)}\Sigma_k B_{ll'k}\, G_k(nl,n'l')\right] \qquad (10)$$

Now $$I(nl) = -\frac{1}{2}\int P(nl)\left[\frac{d^2}{dr^2} + \frac{2N}{r} - \frac{l(l+1)}{r^2}\right]P(nl)\,dr$$

and substituting for the second factor in the integrand from (2), (3), (6), (7) and using the property of orthogonality between the (nl) outer wave function and the core wave functions of the same l, we find

$$2E_1 = -\epsilon_{nl,nl} \qquad (11)$$

Thus if the series electron wave functions are calculated from Fock's equation, and made orthogonal to the appropriate core wave functions, the diagonal parameters give directly the negative energy values in Rydbergs (since 1 Rydberg = ½ atomic unit of energy). This is a result previously quoted (D. R. and W. Hartree 1936*a*, p. 605) without proof.

In Table V the values of $\epsilon_{nl,nl}$ are compared both with the observed values of ν/R and with the values of $\epsilon_{nl,nl}$ for the self-consistent field without exchange; in the latter case the values of $\epsilon_{nl,nl}$ do not give directly the values of $2E_1$, but the rather elaborate calculations required to obtain values of $2E_1$ from them (cf. McDougall 1932) have not been carried out.

TABLE V—Ca⁺ ENERGIES

	Fock's equation $\epsilon_{nl,nl}$	Without exchange $\epsilon_{nl,nl}$	Obs. ν/R
4*s*	0.829$_5$	0.763	0.872$_5$
4*p*	0.619	0.569	0.642
3*d*	0.666	0.507	0.748

The change in the value of $\epsilon_{3d,\,3d}$ brought about by the inclusion of exchange terms is considerable, and suggests that exchange effects have a considerable part in determining the value of N for which the binding of a (3*d*) electron becomes stronger than that of a (4*s*), and so in determining the position of the beginning of the group of transition elements in the periodic table

7—Ca$^+$ TRANSITION PROBABILITIES

Approximate calculation of transition probabilities for the (4*p*)–(3*d*) and (4*p*)–(4*s*) transitions of Ca$^+$ were made some years ago by Zwaan (1929), using a form of Jeffreys (1924) method for the approximate solution of the wave equation and taking an empirical field adjusted to make the ϵ parameters of the one-electron wave functions in this field agree with the observed energy levels of the optical and X-ray spectra McDougall (1932, footnote on p 575) showed that this use of an empirical field to reproduce the observed energy levels of the optical spectrum would not necessarily give the best wave functions for the series electron, and that wave functions calculated, even without exchange, in the self-consistent field of the core would in many cases provide a better approximation Also Zwaan's wave functions (1929, Table VI) do not appear to have the right behaviour near the origin, and the method of calculating them was rather rough Hence a revision of the calculated transition probabilities, using the best available wave functions, seems to be called for

The matrix elements of r, calculated (*a*) from the solution of Fock's equation, (*b*) from the wave functions without exchange in the self-consistent field of the core, and (*c*) by Zwaan, are given in Table VI

TABLE VI

	(*a*)	(*b*)	(*c*)	
$\int_0^\infty rP(4s)\,P(4p)\,dr$	−3·93	−4·28	−3·80	atomic units
$\int_0^\infty rP(4p)\,P(4d)\,dr$	2·63	4·08	2·50	

In order to obtain the observed energy values for the optical terms, Zwaan had to use a field which, for large r, is considerably greater than the self-consistent field of the core This difference might be and indeed used to be, ascribed to the polarization of the core by the series electron, but it is now recognized that exchange effects between the outer wave function and the core wave functions make appreciable contribution to the energy values for the series electron (cf McDougall 1932), and that therefore the energy values of a single electron in a static field should certainly not correspond exactly with the energy levels of the actual atom However, for the wave function of the series electron, the exchange terms in the equations from which they are derived have qualitatively the same effect as an additional attractive field, but the rather close quantitative agreement between Zwaan's values in Table VI and those calculated from the wave functions of this paper appears accidental

It is convenient to use the concept of line strengths introduced by Condon and Shortley (1935, Chap 4, § 7) as an intermediate step in the calculation of transition probabilities The total strength of the multiplet (cf Condon and Shortley 1935, Chap 9, § 2) arising from the transition (nl)–$(n'l')$ in a one-electron spectrum is defined as

$$\mathsf{S}(nl, n'l') = \sum_{\substack{m_l, m_l' \\ m_s, m_s'}} \left| \int \psi^*(nl\, m_l m_s)\, \mathbf{r}\, \psi(n'l'm_l'm_s')\, d\tau \right|^2$$

Now using Sugiura's result (Sugiura 1927, p 113, formula following (4)),

$$\sum_{m_l, m_l'} \left| \int \psi^*(nl m_l m_s)\, \mathbf{r}\, \psi(n'l'm_l'm_s')\, d\tau \right|^2$$
$$= [\mathrm{Max}(l, l')] \left| \int_0^\infty r P(nl)\, P(n'l')\, dr \right|^2 \quad (l' = l \pm 1)$$

(the radial wave functions being normalized), and the sum over m_s, m_s' introduces a factor 2, hence

$$\mathsf{S}(nl, n'l') = 2[\mathrm{Max}(l, l')] \left| \int_0^\infty r P(nl)\, P(n'l')\, dr \right|^2 \tag{12}$$

This result can also be obtained as a special case of formula (3) of Condon and Shortley (1935, Chap 9, § 3)

To convert this result into formulae for oscillator strengths f or transition probabilities A it is necessary to specify which state is the uppermost We will take $(n'l')$ to be the upper and (nl) the lower state Then the total oscillator strength for the transition is

$$f(n'l', nl) = \frac{1}{3} \frac{\nu}{R} \frac{\mathsf{S}(n'l', nl)}{2(2l+1)}$$

(see Condon and Shortley 1935, Chap 5, § 9, formula (1)), where ν is the frequency associated with the transition, and R is the Rydberg frequency, whence from (12)

$$f(n'l', nl) = \frac{1}{3}\frac{\nu}{R}\frac{\text{Max}(l, l')}{2l+1}\left|\int_0^\infty rP(nl)\,P(n'l')\,dr\right|^2 \qquad (13)$$

The oscillator strength f is a pure number, and in this result it is assumed that S and **r** are expressed in atomic units

The statistical weight of the upper $(n'l')$ state is $2(2l'+1)$, and hence, following the argument leading to formula (3) of Condon and Shortley (1935, Chap 4, § 7), we have for the spontaneous transition probability

$$\mathsf{A}(n'l', nl) = \frac{1}{2(2l'+1)}\frac{64\pi^4\nu^3}{3hc^3}[(ea)^2\mathsf{S}(n'l', nl)],$$

if, as in formula (14), S is expressed in atomic units $(ea)^2$ Now

$$\frac{64\pi^4R^3}{3hc^3}(ea)^2 = 2\ 66\ \ 10^9\ \text{sec}^{-1}$$

(Condon and Shortley, 1935, Chap 5, § 6, formula (5)), hence

$$\mathsf{A}(n'l', nl) = 2\ 66\ \ 10^9\frac{\text{Max}(l, l')}{2l'+1}\left(\frac{\nu}{R}\right)^3\left|\int_0^\infty rP(nl)\,P(n'l')\,dr\right|^2 \qquad (14)$$

In calculating values of f and A from (13) and (14), one could use either observed values of ν for the transitions, or calculated values derived from the calculated energy levels The calculated values of ν and of

$$\int_0^\infty rP(nl)\,P(n'l')\,dr$$

are both based on an approximation, and, as far as we know, there is no reason to expect the effects of the approximation on the $(\nu/R)^3$ factor to compensate for the effects on the $\left|\int_0^\infty rP(nl)\,P(n'l')\,dr\right|^2$ factor, hence it seems best to use the observed values of ν/R

Values of f and A for the two transitions $(4p)$–$(4s)$ and $(4p)$–$(3d)$ calculated from the results of this paper using the observed values ν/R are given in Table VII, the values of ν/R derived from the energy levels calculated with exchange are also given

TABLE VII

Transition	l	l'	Obs ν/R	Calc ν/R	S	f	A
$(4p)$–$(4s)$	0	1	$0\ 230_5$	$0\ 210_5$	30 9	1 19	$1\ 66\ \ 10^8\ \text{sec}^{-1}$
$(4p)$–$(3d)$	2	1	0 106	0 047	27 7	0 098	$1\ 46\ \ 10^7\ \text{sec}^{-1}$

8—NEUTRAL Ca WAVE FUNCTIONS AND ENERGIES

The wave functions for the three states of neutral Ca for which calculations have been made are given in Table VIII, the third decimal in the

TABLE VIII—NEUTRAL Ca NORMALIZED WAVE FUNCTIONS

Table of P_N/r^{l+1}

r	$(4s)^2$ state	$(4s)\,(4p)\,^3$P state		$(4s)\,(4p)\,^1$P state	
	$(4s)$	$(4s)$	$(4p)$	$(4s)$	$(4p)$
0 00	4 11	4 28	16 27	4 92	7 23
0 01	3 36	3 49	14 73	4 01	6 55
0 02	2 70	2 80	13 34	3 23	5 93
0 03	2 14	2 21	12 09	2 56	5 38
0 04	1 67	1 72	10 97	2 00	4 88
		Table of P_N			
0 00	0 000	0 000	0 000	0 000	0 000
0 01	0 034	0 035	0 001$_5$	0 040	0 000$_7$
0 02	0 054	0 056	0 005	0 065	0 002$_4$
0 03	0 064	0 066	0 011	0 077	0 004$_8$
0 04	0 067	0 059	0 017$_5$	0 080	0 007$_8$
0 05	0 064	0 065	0 025	0 076	0 011
0 06	0 056	0 057	0 033	0 066	0 014
0 07	0 045	0 046	0 040	0 053	0 018
0 08	0 032	0 033	0 048	0 038	0 021
0 10	0 003	0 003	0 061	0 004	0 027
0 12	−0 027	−0 028	0 073	−0 031	0 032$_5$
0 14	−0 054	−0 056	0 082	−0 064	0 036
0 16	−0 077	−0 079	0 087	−0 092	0 039
0 18	−0 096	−0 098	0 090	−0 114	0 040
0 20	−0 109	−0 111	0 091	−0 130	0 040$_5$
0 22	−0 117	−0 120	0 089	−0 140	0 039$_5$
0 24	−0 120	−0 123	0 085	−0 144	0 038
0 26	−0 120	−0 123	0 079	−0 143	0 035
0 28	−0 115	−0 118	0 072	−0 137	0 031$_5$
0 30	−0 107	−0 110	0 063	−0 128	0 028
0 35	−0 076	−0 078	0 037	−0 091	0 016
0 40	−0 035	−0 037	0 007	−0 043	0 003
0 45	+0 009	+0 009	−0 024	+0 011	−0 011
0 50	0 054	0 055	−0 053	0 063	−0 024
0 55	0 095	0 097	−0 080	0 113	−0 036
0 60	0 131	0 134	−0 104	0 156	−0 046$_5$
0 7	0 184	0 188	−0 141	0 218	−0 063
0 8	0 210	0 215	−0 162	0 249	−0 072$_5$
0 9	0 212	0 217	−0 169	0 251	−0 076
1 0	0 194	0 199	−0 165	0 229	−0 074$_5$
1 1	0 163	0 167	−0 151	0 191	−0 068$_5$
1 2	0 121	0 124	−0 130	0 141	−0 060

TABLE VIII—*continued*

Table of P_N

r	$(4s)^2$ state	$(4s)(4p)\ ^3$P state		$(4s)(4p)\ ^1$P state	
	$(4s)$	$(4s)$	$(4p)$	$(4s)$	$(4p)$
1 4	0 022	0 022	-0 074	0 021	-0 036
1 6	-0 082	-0 086	-0 008	-0 105	$-0\ 007_5$
1 8	-0 182	-0 189	+0 061	-0 224	$+0\ 022_5$
2 0	-0 270	-0 281	0 130	-0 327	0 053
2 2	-0 345	-0 358	0 194	-0 413	0 082
2 4	-0 406	-0 421	0 252	-0 480	$0\ 109_5$
2 6	-0 453	-0 469	0 305	-0 531	0 136
2 8	-0 488	-0 506	0 350	-0 564	0 161
3 0	-0 513	-0 531	0 388	-0 584	0 184
3 2	-0 529	-0 545	0 419	-0 591	0 206
3 4	-0 535	-0 550	0 444	-0 588	0 226
3 6	-0 535	-0 548	0 462	-0 576	0 245
3 8	-0 529	-0 541	0 476	-0 558	0 263
4 0	-0 518	-0 528	0 484	-0 535	0 279
4 5	-0 478	-0 481	0 485	-0 464	0 314
5 0	-0 427	-0 422	0 466	-0 385	0 342
5 5	-0 372	-0 361	0 435	-0 310	0 362
6 0	-0 318	-0 302	0 396	-0 242	0 373
6 5	-0 268	-0 249	0 353	-0 185	0 379
7 0	-0 223	-0 202	0 311	-0 139	0 377
8	-0 150	-0 129	0 232	-0 074	0 358
9	-0 099	-0 080	0 168	-0 036	0 325
10	-0 064	-0 048	0 117	-0 016	0 285
11	-0 040	-0 029	0 081	-0 006	0 242
12	-0 025	-0 017	0 055	-0 002	0 201
14	-0 009	-0 006	0 025	—	$0\ 130_5$
16	-0 003	-0 002	0 011	—	0 080
18	-0 001	-0 001	0 005	—	0 047
20	—	—	0 002	—	0 027
22	—	—	0 001	—	0 015
24	—	—	—	—	0 008
26	—	—	—	—	0 004
28	—	—	—	—	0 002
30	—	—	—	—	0 001
32	—	—	—	—	$0\ 000_5$
$\epsilon_{nl,\,nl}$	0 3891	0 5177	0 3058	0 5052	0 1720
		$\epsilon_{3s\ 4s}$ $= -0\ 007_3$	$\epsilon_{3p\ 4p}$ $= -0\ 013_8$		

tabulated values is not everywhere certain to ± 1 but is probably reliable to ± 2 The wave functions for the $(4s)(4p)$ configuration are shown graphically in fig 2 Comparison with the figure showing the corresponding wave functions for Be (D R and W Hartree 1936*a*, fig 1) shows that the general effect of exchange terms in the two cases is very similar

The energies E_s of the states of neutral Ca relative to the normal state of Ca^{++} as zero, can be written down from the expressions for the corresponding states of Be (D R and W Hartree 1936*a*, formulae (33), (34) and Table II) by appropriate changes of the principal quantum numbers The energies E_s, and contributions to them, are shown in Table IX, together with the corresponding energies derived from observed term values The energies relative to the normal state of Ca^{+} are also shown, for the calculated values, the value $2E_1 = -0\ 829_5$ for Ca^{+}, found from the solution of Fock's equation, has been used

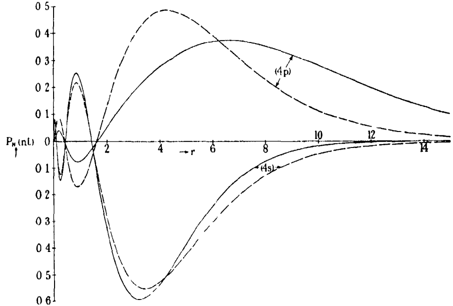


FIG 2—Normalized radial wave functions for the (4*s*) (4*p*) ^{1}P (full line) and ^{3}P (broken line) states of neutral Ca

TABLE IX—NEUTRAL Ca ENERGIES

(All energies are expressed in Rydbergs)

$(4s)^2\ ^1S$		$(4s)\ (4p)\ ^3P$		$(4s)\ (4p)\ ^1P$	
$-2\epsilon_{4s,\,4s}$	$= -0\ 778$	$-\epsilon_{4s\ 4s}$	$= -0\ 518$	$-\epsilon_{4s,\,4s}$	$= -0\ 505$
$-2F_0(4s, 4s)$	$= -0\ 428$	$-\epsilon_{4p\ 4p}$	$= -0\ 306$	$-\epsilon_{4p\ 4p}$	$= -0\ 172$
		$-2F_0(4s, 4p)$	$= -0\ 390$	$-2F_0(4s, 4p)$	$= -0\ 288$
		$+\frac{1}{3}G_1(4s, 4p)$	$= +0\ 085$	$-\frac{1}{3}G_1(4s, 4p)$	$= -0\ 032$
Total $= 2E_s$	$-1\ 206$		$-1\ 129$		$-0\ 997$
Obs $2E_s$	$-1\ 322$		$-1\ 184$		$-1\ 106$
Calc $2E_s - 2E_1(4s)$	$-0\ 376_5$		$-0\ 299_5$		$-0\ 167_5$
Obs $2E_s - 2E_1(4s)$	$-0\ 449$		$-0\ 311$		$-0\ 234$

The agreement between calculated and observed energies is better for the ^{3}P than for the ^{1}P state, for Be, the agreement was better for the ^{1}P state, but the two cases are not closely comparable, as for Be the $(2p)$ wave function corresponds to a non-penetrating orbit, and its exchange interaction with the core is very small, whereas for Ca the $(4p)$ wave function corresponds to a penetrating orbit and its exchange interaction with the core is much greater

9—NEUTRAL Ca TRANSITION PROBABILITIES

As already mentioned, the ^{1}S–^{1}P transition gives an astrophysically important line, and the main object of the calculating of the wave functions for the ^{1}P state was for the evaluation of this transition probability This transition is a "one electron jump" in the sense that the quantum numbers of only one of the wave functions change, but this change is accompanied by an alteration of the other outer wave function, and this should be taken into account in evaluating the transition probability

Writing A for the configuration $(4s)^2$ and B for the configuration $(4s, 4p)$ and, $\psi_X(nl\,m_l m_s \,|\, j)$ $(X = A$ or $B)$ for the $(nl\,m_l m_s)$ wave function in configuration X, regarded as a function of the coordinates of electron j, and also

$$\Phi(A) = \begin{vmatrix} \psi_A(4s0\tfrac{1}{2} \,|\, 1) & \psi_A(4s0-\tfrac{1}{2} \,|\, 1) \\ \psi_A(4s0\tfrac{1}{2} \,|\, 2) & \psi_A(4s0-\tfrac{1}{2} \,|\, 2) \end{vmatrix}$$

$$\Phi_1(B) = \begin{vmatrix} \psi_B(4s0\tfrac{1}{2} \,|\, 1) & \psi_B(4p0-\tfrac{1}{2} \,|\, 1) \\ \psi_B(4s0\tfrac{1}{2} \,|\, 2) & \psi_B(4p0-\tfrac{1}{2} \,|\, 2) \end{vmatrix},$$

$$\Phi_2(B) = \begin{vmatrix} \psi_B(4s0-\tfrac{1}{2} \,|\, 1) & \psi_B(4p0\tfrac{1}{2} \,|\, 1) \\ \psi_B(4s0-\tfrac{1}{2} \,|\, 2) & \psi_B(4p0\tfrac{1}{2} \,|\, 2) \end{vmatrix}$$

the normalized wave functions for A and for the $(^1\text{P}, M_L = 0)$ state of B are

$$\Psi(A) = \frac{1}{\sqrt{2}}\Phi(A), \quad \Psi(B) = \frac{1}{2}[\Phi_1(B) - \Phi_2(B)],$$

so that the matrix element of $\mathbf{r}_1 + \mathbf{r}_2$ between the two states is

$$\frac{1}{2\sqrt{2}} \sum_{\sigma_1\sigma_2} \int \Phi(A)(\mathbf{r}_1 + \mathbf{r}_2)\,[\Phi_1(B) - \Phi_2(B)]\, d\tau_1\, d\tau_2 \tag{15}$$

Summation over the spin coordinates σ gives

$$\begin{aligned}\sum_{\sigma_1\sigma_2}\Phi(A)\Phi_1(B) &= -\sum_{\sigma_1\sigma_2}\Phi(A)\Phi_2(B)\\ &= \psi_A(4s\mid 1)\,\psi_B(4s\mid 1)\,\psi_A(4s\mid 2)\,\psi_B(4p\mid 2)\\ &\quad+\psi_A(4s\mid 1)\,\psi_B(4p\mid 1)\,\psi_A(4s\mid 2)\,\psi_B(4s\mid 2)\qquad(16)\end{aligned}$$

On multiplication by $\mathbf{r}_1+\mathbf{r}_2$ and integration over τ_1, τ_2 each term of (16) gives a contribution $\left[\int\psi_A(4s)\,\psi_B(4s)\,d\tau\right]\left[\int\psi_A(4s)\,\mathbf{r}\psi_B(4p)\,d\tau\right]$ and in (15) there are four such terms altogether, hence (15) becomes

$$\sqrt{2}\left[\int\psi_A(4s)\,\psi_B(4s)\,d\tau\right]\left[\int\psi_A(4s)\,\mathbf{r}\,\psi_B(4p)\,d\tau\right]\qquad(17)$$

This differs from the matrix element for the transition of one electron in a static field by the factor $\int\psi_A(4s)\,\psi_B(4s)\,d\tau$, whose difference from unity expresses the effect of the alteration of the (4s) wave function on the probability of transition of the electron whose quantum numbers change, and by the factor $\sqrt{2}$ The summation of the squares of the matrix elements for different values of M_L follows exactly the corresponding calculation for transition of a single electron in a static field, but there is now no summation over M_S, whereas for a single electron in a static field this introduced a factor 2 Hence the total line strength of the transition is

$$\mathsf{S}[(4s)^2\,{}^1\mathrm{S},\ (4s)\,(4p)\,{}^1\mathrm{P}] = 2\left[\int P_A(4s)\,P_B(4s)\,dr\right]^2\left[\int rP_A(4s)\,P_B(4p)\,dr\right]^2\qquad(18)$$

as for a single electron

If the alteration of the (4s) wave function is disregarded this gives

$$\mathsf{S}[(4s)^2\,{}^1\mathrm{S},\ (4s)\,(4p)\,{}^1\mathrm{P}] = 2\left[\int rP_A(4s)\,P_B(4p)\,dr\right]^2$$

This is in agreement with Condon and Shortley (1935), Chap 9, § 3, since for the transition between $(4s)(ns)\,{}^1\mathrm{S}$ and $(4s)(n'p)\,{}^1\mathrm{P}$ ($n>4$ so that the electron making the transition is not equivalent to any other in either configuration) formulae (5), (6) of Condon and Shortley (1935, Chap 9 § 3) apply, and give

$$\mathsf{S}[(4s)\,(ns)\,{}^1\mathrm{S},\ (4s)\,(np)\,{}^1\mathrm{P}] = \left[\int rP(ns)\,P(n'p)\,dr\right]^2\qquad(n>4),$$

and if one configuration contains equivalent electrons the absolute strengths of the multiplets arising from transition between unexcluded states are twice as large as for the same transition of the jumping electron when it is not equivalent to another electron in either configuration (Condon and Shortley 1935, p 247)

Since for the transition $(4s)^2$–$(4s)(4p)$ the statistical weight of the upper state, $(2S+1)(2L+1)$, is 3, the transition probability is

$$\mathsf{A}[(4s)(4p)\,{}^1\mathrm{P},\ (4s)^2\,{}^1\mathrm{S}] = 2\ 66\ \ 10^9\ \frac{1}{3}\left(\frac{\nu}{R}\right)^3 \mathsf{S}\ \mathrm{sec}^{-1},$$

where S is in atomic units and is given by (20) The relevant quantities for the calculation of this transition probability are

$$\int P_A(4s)\,P_B(4s)\,dr = 0\ 981$$

$$\int rP_A(4s)\,P_B(4p)\,dr = -4\ 06,$$

$$\nu/R\,(\mathrm{obs}\,) = 0\ 215,$$

$$\mathsf{S}[(4s)(4p)\,{}^1\mathrm{P},\ (4s)^2\,{}^1\mathrm{S}] = 31\ 9,$$

whence $$\mathsf{A}[(4s)(4p)\,{}^1\mathrm{P},\ (4s)^2\,{}^1\mathrm{S}] = 2\ 81\ \ 10^8\,\mathrm{sec}^{-1}$$

The ratio of the calculated $(4p)$–$(4s)$ transition probabilities for Ca and Ca^+ is 1 69 Although the data from which both transition probabilities are evaluated are based on an approximation to the form of the wave function for a many-electron atom, it may be expected that the effects of this approximation will be much the same for the two states of ionization, so that this ratio, which for some applications is more important than the individual transition probabilities, is not likely to be seriously affected by this approximation

The calculation of the transition probability for the intercombination line ${}^1\mathrm{S}$–${}^3\mathrm{P}$ is more elaborate, and probably less certain also, as it involves consideration of the departures from Russell-Saunders coupling and the quantitative introduction of spin forces, and it will not be further considered here The following are values of integrals which may be required in this contest (the wave functions $P_c(4s)$, $P_c(4p)$ are those calculated from Fock's equations for the ${}^3\mathrm{P}$ state)

$$\int P_A(4s)\,P_c(4s)\,dr = 0\ 998,$$

$$\int rP_A(4s)\,P_c(4p)\,dr = -4\ 49$$

10—SUMMARY

Calculations of the self-consistent field, with exchange, have been made for Ca^{++}, and the results are given

Solutions of Fock's equations, without allowance for the perturbation of the core by the series electrons, have also been carried out for the $(4s)$, $(4p)$ and $(3d)$ states of Ca^{+} and $(4s)^2$, $(4s)(4p)$ ^{3}P and ^{1}P states of neutral Ca Energy values for all these states have been calculated and are compared with the observed values, it appears that exchange effects have considerable influence on the relative strengths of binding of a $(4s)$ and a $(3d)$ electron, and so are significant from the point of view of the position of the beginning of the transition group of elements in the periodic table

Transition probabilities have also been calculated for the $(4s)$–$(4p)$ and $(3d)$–$(4p)$ transitions of Ca^{+} and the $(4s)^2$–$(4s)(4p)$ singlet transition of neutral Ca, these transitions all give lines of astrophysical interest, and absolute values of transition probabilities may have astrophysical applications

REFERENCES

Condon, E U and Shortley, G H 1935 "The Theory of Atomic Spectra" Cambridge
Fock, V 1930 *Z Phys* **61**, 126–48
Fock, V and Petrashen, M J 1934 *Phys Z Sowjet* **6**, 368–414
— — 1935 *Phys Z Sowjet* **8**, 547–61
Hartree, D R and W 1935 *a* *Proc Roy Soc* A, **149**, 210–31
— — 1935 *b* *Proc Roy Soc* A, **150**, 9–33
— — 1936 *a* *Proc Roy Soc* A, **154**, 588–607
— — 1936 *b* *Proc Roy Soc* A, **156**, 45–62
— — 1936 *c* *Proc Roy Soc* A **157**, 490–502
Jeffreys, H 1924 *Proc Lond Math Soc* **23**, 428–36
McDougall, J 1932 *Proc Roy Soc* A, **138**, 550–79
Slater, J C 1929 *Phys Rev* **34**, 1293–1322
Sugiura, Y 1927 *J Phys Radium*, series VI, **8**, 118–24
Torrance, C C 1934 *Phys Rev* **46**, 388–90
Zwaan, A 1929 Intensitaten in Ca Funkenspectrum Dissertation, Utrecht

On the Statistical Theory of Isotropic Turbulence

BY THEODORE DE KÁRMÁN AND LESLIE HOWARTH*

California Institute of Technology, Pasadena

(*Communicated by G I Taylor, F R S —Received* 27 *January* 1937—
Revised 6 *November* 1937)

INTRODUCTION AND SUMMARY

G I Taylor (1935), in a paper of fundamental importance, introduced the conception of isotropic turbulence and applied it, with interesting results, to the problem of the decay of turbulence in a windstream

In this paper we develop a general theory of isotropic turbulence The correlation coefficients between two arbitrary velocity components at two arbitrary points form a tensor called the correlation tensor Due to isotropy the tensor corresponding to one fixed and one variable point has spherical symmetry, due to the condition of continuity it is completely determined by one scalar function The mean products of the derivatives of the velocity fluctuations are expressed by the derivatives of the tensor components, in this way laborious calculations for obtaining such mean values are eliminated The correlation between three components (triple correlation) is discussed

After developing the kinematics of isotropic turbulence the dynamical problem of the change of the various mean values with time is considered It is shown that using the equations of motion a partial differential equation connecting the double and triple correlation functions can be established The solution of this equation is investigated first in the case when the triple correlation is neglected and the shape of the double correlation function remains similar and only its scale changes Equations for the dissipation of energy and vorticity are deduced

Finally, a possible solution for large Reynolds numbers is given and applied to Taylor's problem of the decay of turbulence behind a grid

Taylor's fundamental relation between the width of the correlation function and the size of the small (dissipative) eddies is confirmed However, it is believed that the linear law for the reciprocal of the root mean square of the velocity fluctuations as a function of the distance from the grid is a special case, in the last section of the present paper a more general functional relation is suggested

---

* Research Fellow, King's College, Cambridge

§§ 6, 8 and 9 have been rewritten and § 11 inserted by the senior author in September 1937

A—THE KINEMATICS OF ISOTROPIC TURBULENCE

1—*The definition of isotropic turbulence and immediate deductions*

Isotropic turbulence may be defined by the condition that the average value of any function of the velocity components and their derivatives at a particular point, defined in relation to a particular set of axes, is unaltered if the axes of reference are rotated in any manner and if the co-ordinate system is reflected in any plane through the origin We consider average values with regard to the time and suppose that the fluctuations are so rapid that the variation of the average value is negligible throughout the period of time required for averaging Thus, in fact, we shall consider the average values to be slowly varying functions of the time

Consider a particular system of co-ordinate axes Ox_1, Ox_2 and Ox_3 and take two points $P(x_1, 0, 0)$ and $P'(x_1', 0, 0)$ on Ox_1 Denote by u_1, u_2, u_3 and u_1', u_2', u_3' the velocity components at P and P' respectively Let us now suppose that $\overline{u_1^2}$, $\overline{u_2^2}$, $\overline{u_3^2}$ are all independent of position (By isotropy, of course, $\overline{u_1^2} = \overline{u_2^2} = \overline{u_3^2}$ ($= \overline{u^2}$, say))

The correlation coefficients $\overline{u_1u_1'}/\overline{u^2}$ and $\overline{u_ju_j'}/\overline{u^2}$ for $j = 2$ or 3 will be particular functions $f(r, t)$ and $g(r, t)$, say, of the distance r between P and P'* and of the time It seems to be fairly evident physically that the mean value $\overline{u_iu_j'} = 0$ when $i \neq j$ in isotropic turbulence, since it appears to be equally probable that u_1u_2', for example, will be positive or negative It is an easy matter to prove these results analytically The mean values $\overline{u_1u_j'}$ and $\overline{u_1'u_j}$ for $j = 2$ or 3 can be shown to vanish by a rotation of the axes of x_2 and x_3 through 180° about the x_1-axis Denoting transformed values by capital letters we see that $U_1 = u_1$, $U_1' = u_1'$, $U_j = -u_j$ and $U_j' = -u_j'$ for $j = 2$ or 3 so that, for example,

$$\overline{U_1U_j'} = -\overline{u_1u_j'}$$

But by the isotropic property

$$\overline{U_1U_j'} = \overline{u_1u_j'},$$

---

* It is assumed that the turbulence is statistically uniform as well as isotropic so that correlations do not depend on the position or orientation of the line PP but only on its length

so that $\overline{u_1 u_j'}$ is zero for $j = 2$ or 3 Similarly $\overline{u_j u_1'}$ is also zero for $j = 2$ or 3 The mean values $\overline{u_2 u_3'}$ and $\overline{u_3 u_2'}$ can be shown to vanish by reflexion in the x_1, x_3 plane Denoting, in this case, reflected values by capital letters we see that $U_2 = -u_2$, $U_2' = -u_2'$, $U_3 = u_3$, $U_3' = u_3'$, so that, for example,

$$\overline{U_2 U_3'} = -\overline{u_2 u_3'}$$

But by the isotropic property

$$\overline{U_2 U_3'} = \overline{u_2 u_3'},$$

and hence $\overline{u_2 u_3'}$ vanishes Similarly $\overline{u_3 u_2'}$ is zero

2—*Lemma*

Consider now any two points P and Q in the fluid Denote by p the velocity component in a particular direction PP' at P and by q the velocity component in the direction QQ' at Q We shall determine an expression for the correlation coefficient $\overline{pq}/\overline{u^2}$

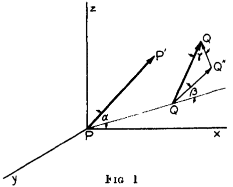


FIG 1

The intersecting lines PP' and PQ (see fig 1) determine a plane Denote by QQ'' the orthogonal projection of QQ' on this plane, $Q''Q'$ will then be normal to the plane Denote by α, β and γ the angles $P'\hat{P}Q$, $\pi - P\hat{Q}Q''$, $Q\hat{Q}'Q''$, and by p_1, p_2, p_3 and q_1, q_2, q_3, respectively, the velocity components at P and Q in the direction PQ, the direction normal to PQ and $Q''Q'$, and the direction $Q''Q'$ * Then

$$p = p_1 \cos\alpha + p_2 \sin\alpha,$$
$$q = q_1 \cos\beta \sin\gamma + q_2 \sin\beta \sin\gamma + q_3 \cos\gamma$$

---

* For simplicity the figure is described as it is drawn in fig 1 When P' and Q'' are on opposite sides of PQ the sign of β must be changed

Therefore

$$\overline{pq} = \overline{p_1q_1}\cos\alpha\cos\beta\sin\gamma + \overline{p_1q_2}\cos\alpha\sin\beta\sin\gamma + \overline{p_1q_3}\cos\alpha\cos\gamma$$
$$+\overline{p_2q_1}\cos\beta\sin\gamma\sin\alpha + \overline{p_2q_2}\sin\beta\sin\alpha\sin\gamma + \overline{p_2q_3}\sin\alpha\cos\gamma$$

We have already proved that all the mean values here except $\overline{p_1q_1}$ and $\overline{p_2q_2}$ vanish, and we have denoted by $f(r,t)$ and $g(r,t)$ the correlation coefficients $\overline{p_1q_1}/\overline{u^2}$ and $\overline{p_2q_2}/\overline{u^2}$, where PQ is supposed to be of length r Thus

$$\frac{\overline{pq}}{\overline{u^2}} = [f(r,t)\cos\alpha\cos\beta + g(r,t)\sin\alpha\sin\beta]\sin\gamma \qquad (1)$$

3—*The correlation tensor*

Consider now any particular co-ordinate system and suppose the velocity components at $P(x_1, x_2, x_3)$ and $P'(x_1', x_2', x_3')$ are (u_1, u_2, u_3) and (u_1', u_2', u_3') respectively The nine quantities $\overline{u_iu_j'}$ for $i, j = 1, 2$ or 3 may be shown to be the components of a second rank tensor The transformation law can readily be seen to be satisfied since the dyadic product of two vectors is a tensor so that each of the contributions u_iu_j' to the mean values form a tensor Clearly the operation of taking a mean value will not alter the transformation law satisfied Hence we may consider the "correlation tensor" **R** defined by

$$\overline{u^2}\mathbf{R} = \overline{u^2}\begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} = \begin{pmatrix} \overline{u_1u_1'} & \overline{u_1u_2'} & \overline{u_1u_3'} \\ \overline{u_2u_1'} & \overline{u_2u_2'} & \overline{u_2u_3'} \\ \overline{u_3u_1'} & \overline{u_3u_2'} & \overline{u_3u_3'} \end{pmatrix}$$

or
$$\overline{u^2}\mathbf{R} = \overline{u^2}R_{ij} = \overline{u_iu_j'} \qquad i,j = 1,\ 2 \text{ or } 3 \qquad (2)$$

By means of the lemma which has been established each component of **R** can be evaluated in terms of the functions f and g and the vector $\mathbf{r}$ whose components are $\xi_1 = x_1' - x_1$, $\xi_2 = x_2' - x_2$, $\xi_3 = x_3' - x_3$ In this way we obtain

$$\mathbf{R} = \frac{\{f(r,t) - g(r,t)\}}{r^2}\mathbf{rr} + g(r,t)\,\mathbf{I}, \qquad (3)$$

where $r = |\mathbf{r}|$ and $\mathbf{I}$ is the unit tensor $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

By expressing the fact that the velocity fluctuations satisfy the equation of continuity we can now obtain a relation between $f(r, t)$ and $g(r, t)$ Since the velocity fluctuations at P' satisfy the equation of continuity

$$\frac{\partial u_i'}{\partial x_i'} = 0,$$

where the summation convention is in operation Therefore

$$\frac{\partial}{\partial x_i'}(u_i' u_j) = 0, \tag{4}$$

since u_j is independent of x_1', x_2' and x_3' Thus we obtain the result

$$\frac{\partial R_{ji}}{\partial x_i'} = 0, \tag{5}$$

for $j = 1$, 2 or 3 by dividing equation (4) by $\overline{u^2}$ (which is independent of position), by taking mean values and using the definition of R_{ji} (see equation (2)) By the definition of ξ_1, ξ_2, ξ_3 we may write equation (5) as

$$\frac{\partial R_{ji}}{\partial \xi_i} = 0 \tag{6}$$

Using the values of the components of $\mathbf{R}$ given by equation (3), and also the fact that equation (6) is valid for all values of ξ_1, ξ_2 and ξ_3, we find

$$2f(r, t) - 2g(r, t) = -r\frac{\partial f(r, t)}{\partial r} \tag{7}$$

Since both $f(r, t)$ and $g(r, t)$ are even functions of r

$$f(r, t) = 1 + f_0''\frac{r^2}{2!} + f_0^{\text{iv}}\frac{r^4}{4!} + \quad, \tag{8}$$

$$g(r, t) = 1 + g_0''\frac{r^2}{2!} + g_0^{\text{iv}}\frac{r^4}{4!} + \tag{9}$$

From equation (7) we then find by equating coefficients of r^2 and r^4

$$2f_0'' = g_0'', \tag{10}$$

$$3f_0^{\text{iv}} = g_0^{\text{iv}} \tag{11}$$

Further, substituting these expansions for f and g in equation (3) we find, for small values of r, that

$$\mathbf{R} = \left(1 + \frac{g_0''}{2}r^2\right)\mathbf{I} + \left(\frac{f_0'' - g_0''}{2}\right)\mathbf{rr}, \tag{12}$$

neglecting fourth and higher powers of r Using equation (10) we may write equation (12) in the form

$$\mathbf{R} = (1 + f_0'' r^2)\mathbf{I} - \tfrac{1}{2} f_0'' \mathbf{r}\mathbf{r} \tag{13}$$

We notice, in passing, that

$$\left.\begin{aligned} \left(\frac{\partial^2 R_{kl}}{\partial \xi_i \partial \xi_j}\right)_{\xi_1=\xi_2=\xi_3=0} &= f_0'' \quad \text{when} \quad k = l = i = j, \\ &= 2f_0'' \quad \text{when} \quad k = l \neq i = j, \\ &= -\tfrac{1}{2} f_0'' \quad \text{when} \quad k = i \neq l = j \quad \text{or} \quad k = j \neq l = i, \\ &= 0 \text{ otherwise}, \end{aligned}\right\} \tag{14}$$

we shall require these results later

We may, at this stage, point out an analogy which is helpful in creating a physical picture The expression (3) for the correlation tensor is of exactly the same form as that for the stress tensor for a continuous medium when there is spherical symmetry In the analogy $f(r)$ is the principal radial stress at any point, $g(r)$ is any of the principal transverse stresses, and R_{ik} is the component in the k direction of the stress over a plane whose normal is in the i direction Further, the relation between f and g given by continuity in our problem corresponds to the condition for equilibrium in the stress analogy

4—*The correlation coefficients between derivatives of the velocities*

Taylor (1935) had occasion to calculate the various correlation coefficients between first derivatives of the velocity components These calculations were made by transforming axes, using the equation of continuity and applying the definition of isotropy Even for the first derivatives this method is fairly laborious, whilst for higher order derivatives the work involved makes it almost prohibitive, it is hardly possible to decide beforehand whether a particular transformation will lead to a new relation or will merely give one that has already been obtained

The determination of the correlations both quickly and automatically is an interesting application of the correlation tensor Let us take, as an example, the determination of $\overline{\dfrac{\partial u_k}{\partial x_i}\dfrac{\partial u_l}{\partial x_j}}$ Now

$$\frac{\partial}{\partial x_i}\overline{(u_k u_l')} = \overline{u^2}\frac{\partial}{\partial x_i} R_{kl} = -\overline{u^2}\frac{\partial R_{kl}}{\partial \xi_i}, \tag{15}$$

i.e since the velocity fluctuations at P are all differentiable functions of x_i,

$$\overline{\frac{\partial u_k}{\partial x_i}u_l'} = -\overline{u^2}\frac{\partial R_{kl}}{\partial \xi_i} \tag{16}$$

Further
$$\frac{\partial}{\partial x_j'}\left(\overline{\frac{\partial u_k}{\partial x_i}u_l'}\right) = -\overline{u^2}\frac{\partial}{\partial x_j'}\left(\frac{\partial R_{kl}}{\partial \xi_i}\right) = -\overline{u^2}\frac{\partial^2 R_{kl}}{\partial \xi_j\,\partial \xi_i},$$

i e
$$\overline{\frac{\partial u_k}{\partial x_i}\frac{\partial u_l'}{\partial x_j'}} = -\overline{u^2}\frac{\partial^2 R_{kl}}{\partial \xi_i\,\partial \xi_j} \tag{17}$$

If, now, we make P and P' coincide we obtain the result

$$\overline{\frac{\partial u_k}{\partial x_i}\frac{\partial u_l}{\partial x_j}} = -\overline{u^2}\left(\frac{\partial^2 R_{kl}}{\partial \xi_i\,\partial \xi_j}\right)_{\xi_1=\xi_2=\xi_3=0} \tag{18}$$

We obtain the values of $\dfrac{\partial^2 R_{kl}}{\partial \xi_i\,\partial \xi_j}$ when $\xi_1 = \xi_2 = \xi_3 = 0$ immediately from equations (14) Hence, for example, we see that

$$\overline{\left(\frac{\partial u_1}{\partial x_1}\right)^2} = -\overline{u^2}f_0'', \tag{19}$$

$$\overline{\left(\frac{\partial u_1}{\partial x_2}\right)^2} = -2\overline{u^2}f_0'', \tag{20}$$

and that
$$\overline{\frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_2}} = \overline{\frac{\partial u_1}{\partial x_2}\frac{\partial u_2}{\partial x_1}} = \tfrac{1}{2}\overline{u^2}f_0''$$

$$= -\frac{1}{2}\overline{\left(\frac{\partial u_1}{\partial x_1}\right)^2}, \tag{21}$$

or, alternatively,
$$= -\frac{1}{4}\overline{\left(\frac{\partial u_1}{\partial x_2}\right)^2}, \tag{22}$$

by equations (19) and (20) respectively

Similarly, all other correlations of first derivatives can be obtained

This method is immediately applicable to higher order derivatives Exactly as for the first derivatives we obtain, for example, the result

$$\overline{\frac{\partial^2 u_k}{\partial x_i\,\partial x_j}\frac{\partial^2 u_l}{\partial x_r\,\partial x_s}} = \overline{u^2}\left(\frac{\partial^4 R_{kl}}{\partial \xi_i\,\partial \xi_j\,\partial \xi_r\,\partial \xi_s}\right)_{\xi_1=\xi_2=\xi_3=0} \tag{23}$$

The fourth derivative of the components of **R** can be obtained from equation (3), and using the expansions (8) and (9) we find, for example,

$$\overline{\left(\frac{\partial^2 u_1}{\partial x_1^2}\right)^2} = \overline{u^2}f_0^{\text{iv}}, \tag{24}$$

$$\overline{\left(\frac{\partial^2 u_1}{\partial x_2^2}\right)^2} = 3\overline{u^2} f_0^{\text{iv}}, \tag{25}$$

and

$$\overline{u^2}\left(\frac{\partial^4 R_{11}}{\partial \xi_1^2 \partial \xi_2^2}\right)_{\xi_1=\xi_2=\xi_3=0} = \tfrac{2}{3}\overline{u^2} f_0^{\text{iv}},$$

so that

$$\overline{\frac{\partial^2 u_1}{\partial x_1^2}\frac{\partial^2 u_1}{\partial x_2^2}} = \overline{\left(\frac{\partial^2 u_1}{\partial x_1 \partial x_2}\right)^2} = \frac{2}{9}\overline{\left(\frac{\partial^2 u_1}{\partial x_2^2}\right)^2} = \frac{2}{3}\overline{\left(\frac{\partial^2 u_1}{\partial x_1^2}\right)^2}, \tag{26}$$

and so on

For completeness we include the results for first and second derivatives, writing (u, v, w) and (x, y, z) for (u_1, u_2, u_3) and (x_1, x_2, x_3) For the first derivatives*

$$\overline{\left(\frac{\partial u}{\partial x}\right)^2} = \frac{1}{2}\overline{\left(\frac{\partial u}{\partial y}\right)^2}, \quad \overline{\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}} = \overline{\frac{\partial u}{\partial y}\frac{\partial v}{\partial x}} = -\frac{1}{4}\overline{\left(\frac{\partial u}{\partial y}\right)^2} \tag{27}$$

For the second derivatives

$$\left.\begin{aligned}
&\overline{\left(\frac{\partial^2 u}{\partial x^2}\right)^2} = \overline{\left(\frac{\partial^2 u}{\partial y \partial z}\right)^2} = \overline{\frac{\partial^2 u}{\partial y^2}\frac{\partial^2 u}{\partial z^2}} = \frac{1}{3}\overline{\left(\frac{\partial^2 u}{\partial y^2}\right)^2}, \\
&\overline{\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2} = \overline{\frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial y^2}} = \frac{2}{9}\overline{\left(\frac{\partial^2 u}{\partial y^2}\right)^2}, \\
&\overline{\frac{\partial^2 u}{\partial x^2}\frac{\partial^2 v}{\partial x \partial y}} = \overline{\frac{\partial^2 u}{\partial y^2}\frac{\partial^2 v}{\partial x \partial y}} = -\frac{1}{6}\overline{\left(\frac{\partial^2 u}{\partial y^2}\right)^2}, \\
&\overline{\frac{\partial^2 u}{\partial x \partial z}\frac{\partial^2 v}{\partial y \partial z}} = \overline{\frac{\partial^2 u}{\partial z^2}\frac{\partial^2 v}{\partial x \partial y}} = \overline{\frac{\partial^2 u}{\partial y \partial z}\frac{\partial^2 v}{\partial z \partial x}} = -\frac{1}{18}\overline{\left(\frac{\partial^2 u}{\partial y^2}\right)^2}
\end{aligned}\right\} \tag{28}$$

All mean values which cannot be brought into one of these forms by cyclic permutation of the letters vanish

It is obvious that the process could be continued and the higher derivatives of R_{ik} give analogous expressions for the mean products of the higher derivatives of the velocity components Taylor's development formula for his correlation function R_y (1935, equation (46)) is a special case of the general formulae obtained in this way His equation, which he effectively found as early as 1921, would be expressed in our notation

$$g(r, t) = 1 - \frac{\overline{\left(\frac{\partial u_1}{\partial x_2}\right)^2}}{\overline{u^2}}\frac{r^2}{2!} + \frac{\overline{\left(\frac{\partial^2 u_1}{\partial x_2^2}\right)^2}}{\overline{u^2}}\frac{r^4}{4!} -$$

An analogous expression can be written for $f(r, t)$

---

* These results were first given by Taylor (1935, p 436)

5—*Expression of mean values by integrals*

It is known that the velocity components in an incompressible fluid can be expressed by a vector and a scalar potential In the present case (indefinitely extended fluid without sources throughout the whole fluid) the scalar potential vanishes and the velocity components can be written (it is convenient to use x, y, z, u, v, w for the co-ordinates and the velocity components in this section)

$$u = \frac{\partial H}{\partial y} - \frac{\partial G}{\partial z}, \quad v = \frac{\partial F}{\partial z} - \frac{\partial H}{\partial x}, \quad w = \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}, \tag{29}^*$$

where

$$F = \frac{1}{4\pi}\iiint \frac{\omega_x'}{r}\,dx'\,dy'\,dz', \quad G = \frac{1}{4\pi}\iiint \frac{\omega_y'}{r}\,dx'\,dy'\,dz',$$
$$H = \frac{1}{4\pi}\iiint \frac{\omega_z'}{r}\,dx'\,dy'\,dz', \tag{30}$$

ω_x', ω_y', ω_z' are the components of vorticity at the point (x', y', z'),

$$r^2 = (x'-x)^2 + (y'-y)^2 + (z'-z)^2,$$

and the integration extends throughout the fluid These results enable us to express correlations between vorticity components and velocity derivatives conveniently as integrals Consider, for example, $\overline{\omega_x \frac{\partial w}{\partial y}}$ Now

$$\frac{\partial w}{\partial y} = \frac{1}{4\pi}\left\{\frac{\partial^2}{\partial y\,\partial x}\iiint \frac{\omega_y'}{r}\,dx'\,dy'\,dz' - \frac{\partial^2}{\partial y^2}\iiint \frac{\omega_x'}{r}\,dx'\,dy'\,dz'\right\}$$
$$= \frac{1}{4\pi}\iiint\left[\frac{3(x'-x)(y'-y)}{r^5}\,\omega_y' - \left\{\frac{3(y'-y)^2}{r^5} - \frac{1}{r^3}\right\}\omega_x'\right]dx'\,dy'\,dz', \tag{31}$$

so that

$$\overline{\omega_x \frac{\partial w}{\partial y}} = \frac{1}{4\pi}\iiint\left[\frac{3(x'-x)(y'-y)}{r^5}\,\overline{\omega_x\,\omega_y'} - \left\{\frac{3(y'-y)^2}{r^5} - \frac{1}{r^3}\right\}\overline{\omega_x\,\omega_x'}\right]dx'\,dy'\,dz' \tag{32}$$

Now we can define a correlation tensor **V** for the vorticity components in precisely the same way as **R** was defined for the velocity components Hence, using an obvious notation,

$$\overline{\omega_x\,\omega_y'} = \overline{\omega_x^2}\,V_{xy}, \quad \overline{\omega_x\,\omega_x'} = \overline{\omega_x^2}\,V_{xx}$$

---

* See Lamb (1932) These results are purely kinematic

and

$$\overline{\omega_x \frac{\partial w}{\partial y}} = \frac{\overline{\omega_x^2}}{4\pi}\iiint\left[\frac{3(x'-x)(y'-y)}{r^5}V_{xy} - \left\{\frac{3(y'-y)^2}{r^5}-\frac{1}{r^3}\right\}V_{xx}\right]dx'\,dy'\,dz' \tag{33}$$

When the turbulence is isotropic this method of approach has no special merit since the required results may be obtained more easily by other methods It is clear that the form for **V** is exactly similar to that for **R** (including the condition arising from continuity) in isotropic flow, using this fact the integral in equation (33) can be evaluated and the result compared with the one that can be deduced immediately from equations (27)

6—*Triple correlations**

We shall now consider the mean values of the product of three-velocity components u_i, u_j, u'_k, where u_i and u_j are the instantaneous values of the i and j components of the velocity observed at an arbitrary point P and u'_k is the k-component of the velocity observed at another arbitrary point P' Let us consider again homogeneous isotropic turbulence Then

$$\overline{u_i^2} = \overline{u_j^2} = \overline{u_k'^2} = \overline{u^2},$$

and we write
$$\overline{u_i u_j u'_k} = (\overline{u^2})^{\frac{3}{2}}\, T_{ijk},$$

where T_{ijk} is a tensor of third rank The tensor T_{ijk} will be a function of ξ_1, ξ_2, ξ_3 where $\xi_1 = x'_1 - x_1$, $\xi_2 = x'_2 - x_2$, $\xi_3 = x'_3 - x_3$, we call it the tensor of the triple correlations In order to find the expression for T_{ijk} as a function of the ξ's, let us assume first that both points P and P' lie on the x-axis Then it is obvious that all quantities $\overline{u_i u_j u'_k}$ belong to one of the following six groups, $\overline{u_1^2 u'_1}$, $\overline{u_1^2 u'_k}$, $\overline{u_1 u_i u'_1}$, $\overline{u_1 u_j u'_k}$, $\overline{u_i u_j u'_1}$ and $\overline{u_i u_j u'_k}$, where i, j, k can be equal to 2 or 3 Due to the assumption of isotropy $\overline{u_1^2 u'_k} = 0$, $\overline{u_1 u_i u'_1} = 0$, and $\overline{u_i u_j u'_k} = 0$, because by reflexion of at least one of the axes x_2 or x_3 these expressions certainly change their sign Furthermore, $\overline{u_1 u_j u'_k}$ and $\overline{u_i u_j u'_1}$ vanish for the same reason unless $j = k$ or $i = j$ Hence only the following mean values remain as possibly different from zero $\overline{u_1^2 u'_1}$, $\overline{u_1 u_2 u'_2}$, $\overline{u_1 u_3 u'_3}$, $\overline{u_2^2 u'_1}$, $\overline{u_3^2 u'_1}$ Obviously because of isotropy $\overline{u_1 u_2 u'_2} = \overline{u_1 u_3 u'_3}$ and $\overline{u_2^2 u'_1} = \overline{u_3^2 u'_1}$ Hence three independent quantities remain different from zero, we put $\overline{u_1^2 u'_1} = (\overline{u^2})^{\frac{3}{2}}\, k(r)$, $\overline{u_1 u_2 u'_2} = \overline{u_1 u_3 u'_3} = (\overline{u^2})^{\frac{3}{2}}\, q(r)$ and $\overline{u_2^2 u'_1} = \overline{u_3^2 u'_1} = (\overline{u^2})^{\frac{3}{2}}\, h(r)$

---

* In order to distinguish between correlations connecting two and three velocity components the correlations treated in the previous sections will be referred to in the following sections as "double correlations"

(Fig 2 represents the double and triple correlation functions which do not vanish because of isotropy)

It will be seen that the series development of the function $k(r)$ for small values of r starts with the term in r^3 For, developing u_1' in a series of powers of the variable $\xi_1 = x_1' - x_1 = r$, we obtain

$$\overline{u_1^2 u_1'} = \overline{u_1^3} + \overline{u_1^2 \frac{\partial u_1}{\partial \xi_1}}\,\xi_1 + \tfrac{1}{2}\,\overline{u_1^2 \frac{\partial^2 u_1}{\partial \xi_1^2}}\,\frac{\xi_1^2}{2} +$$

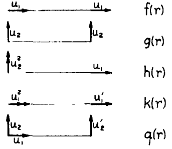


FIG 2

Now, reflecting the x-axis all the coefficients of even powers of ξ_1 change their sign, and consequently these coefficients must vanish in the case of isotropic turbulence Because of the homogeneity of turbulence

$$\overline{u_1^2 \frac{\partial u_1}{\partial \xi_1}} = \frac{1}{3}\frac{\partial \overline{u_1^3}}{\partial \xi_1},$$

also vanishes Hence, the development of $\overline{u_1^2 u_1'}$ starts with the term

$$\tfrac{1}{6}\,\overline{u_1^2 \frac{\partial^3 u_1}{\partial \xi_1^3}}\,\xi_1^3$$

and correspondingly

$$k(r) = k'''(0)\frac{r^3}{6} + k^{\mathrm{v}}(0)\frac{r^5}{120} +$$

Until now it has been assumed that P and P' are situated on the x-axis In order to obtain the general expression for $\overline{u_i u_j u_k'}$ we take $\overline{PP'}$ to be an arbitrary direction and denote by (p_1, p_2, p_3) and (p_1', p_2', p_3') the velocity components at P and P' along three mutually perpendicular lines whose direction cosines are (l_1, m_1, n_1), (l_2, m_2, n_2) and (l_3, m_3, n_3) In particular, we suppose that $l_1 = \xi_1/r$, $m_1 = \xi_2/r$, $n_1 = \xi_3/r$, so that the first line is in the direction $\overline{PP'}$

We then obtain for the velocity components along the x, y, z axes

$$\left.\begin{aligned} u_1 &= l_i p_i, & u_1' &= l_i p_i', \\ u_2 &= m_i p_i, & u_2' &= m_i p_i', \\ u_3 &= n_i p_i, & u_3' &= n_i p_i' \end{aligned}\right\} \tag{34}$$

On the other hand, because of isotropy

$$\overline{p_1^2 p_1'} = (\overline{u^2})^{\frac{3}{2}} k(r),\ \overline{p_1 p_2 p_2'} = \overline{p_1 p_3 p_3'} = (\overline{u^2})^{\frac{3}{2}} q(r) \text{ and } \overline{p_2^2 p_1'} = \overline{p_3^2 p_1'} = (\overline{u^2})^{\frac{3}{2}} h(r)$$

all other combinations being equal to zero

Using the expressions (34) and substituting for the mean values of the products of the p's the functions $k(r)$, $q(r)$ and $h(r)$ we find after some analysis that

$$\overline{u_i u_j u_k'} = (\overline{u^2})^{\frac{3}{2}} T_{ijk} = (\overline{u^2})^{\frac{3}{2}} \left\{\xi_i \xi_j \xi_k \frac{k-h-2q}{r^3} + \delta_{ij} \xi_k \frac{h}{r} + \delta_{ik} \xi_j \frac{q}{r} + \delta_{jk} \xi_i \frac{q}{r}\right\}, \tag{35}$$

where δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$)

It is seen that T_{ijk} is an odd function of the variables ξ_1, ξ_2, ξ_3 By interchanging P and P' we change ξ_1, ξ_2 and ξ_3 into $-\xi_1$, $-\xi_2$ and $-\xi_3$ respectively Hence the mean values of triple products composed of one component measured at P and two components measured at P' can be expressed by the T_{ijk}'s For instance

$$\overline{u_i u_j' u_k'} = -\overline{u_i' u_j u_k} = -\overline{u_k u_j u_i'} = -(\overline{u^2})^{\frac{3}{2}} T_{kji}$$

This relation will be used in § 8

We shall now show that because of the continuity relation between the velocity components the functions k and q can be expressed by h Since u_i and u_j are independent of the variation of x_k', it follows from the continuity equation that

$$\frac{\partial}{\partial x_k'} (\overline{u_i u_j u_k'}) = 0 \tag{36}$$

Hence

$$\frac{\partial}{\partial \xi_k} (\overline{u_i u_j u_k'}) = 0 \quad \text{or} \quad \frac{\partial}{\partial \xi_k} T_{ijk} = 0 \tag{37}$$

By substituting (35) in (37) and carrying out the differentiations, we find

$$\xi_i \xi_j \left[\frac{k'-h'}{r^2} + \frac{2k-2h-6q}{r^3}\right] + \delta_{ij}\left(\frac{2h}{r} + \frac{2q}{r} + h'\right) = 0 \tag{38}$$

The symbols k' and h' denote differential quotients with respect to r The k, h and q being only functions of r, the expressions in the brackets must vanish separately This leads to the results

$$\left.\begin{aligned} k &= -2h, \\ q &= -h - \frac{r}{2}\frac{dh}{dr} \end{aligned}\right\} \qquad (39)$$

It was seen above that the development of $k(r)$ for small values of r starts with the term containing r^3 Because of (39) the same is true for $h(r)$ and $q(r)$

The expression (35) and the relations (39) carry the general analysis of the triple correlations as far as we did in § 3 in the case of the double correlations

7—*The correlation between pressure and velocity*

We shall now prove that the mean values $\overline{\varpi u_j'}$ ($j = 1$, 2 or 3), where ϖ is the pressure at (x_1, x_2, x_3), all vanish Again, denote by (p_1', p_2', p_3') the velocity components at (x_1', x_2', x_3') along the lines whose direction cosines are given by (l_i, m_i, n_i) ($i = 1$, 2 or 3) and defined in the preceding paragraph It is clear, by symmetry, that $\overline{\varpi p_j'}$ ($j = 2$ or 3) vanishes It is, however, necessary to introduce the equation of continuity to show that $\overline{\varpi p_1'} = 0$ Then let us write

$$\varpi p_1' = \{\overline{\varpi^2}\}^{\frac{1}{2}}\{\overline{u^2}\}^{\frac{1}{2}} s(r)$$

We find immediately from (34) that

$$\left.\begin{aligned} \overline{\varpi u_1'} &= \{\overline{\varpi^2}\}^{\frac{1}{2}}\{\overline{u^2}\}^{\frac{1}{2}}[l_1 s(r)], \\ \overline{\varpi u_2'} &= \{\overline{\varpi^2}\}^{\frac{1}{2}}\{\overline{u^2}\}^{\frac{1}{2}}[m_1 s(r)], \\ \overline{\varpi u_3'} &= \{\overline{\varpi^2}\}^{\frac{1}{2}}\{\overline{u^2}\}^{\frac{1}{2}}[n_1 s(r)] \end{aligned}\right\} \qquad (40)$$

Here again, the equation of continuity leads to the condition

$$\frac{\partial}{\partial \xi_i}\overline{(\varpi u_i')} = 0$$

for all values of ξ_1, ξ_2, ξ_3 Transforming to spherical polar co-ordinates we see that

$$\frac{\partial}{\partial r}[r^2 \sin\theta\, s(r)] = 0$$

for all values of r and θ, so that

$$\frac{d}{dr}s(r) + \frac{2s(r)}{r} = 0$$

Again, since s is regular at the origin, the appropriate solution is

$$s = 0$$

Hence, $\overline{\varpi u'_j} = 0$ for $j = 1$, 2 or 3 Similarly, $\overline{\varpi' u_j} = 0$ for $j = 1$, 2 or 3

B—DYNAMICS OF ISOTROPIC TURBULENCE

8—*The equation for the propagation of the correlation*

The equations of motion at the point $P(x_1, x_2, x_3)$ can be written

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \varpi}{\partial x_i} + \nu \nabla_x^2 u_i \tag{41}$$

for $i = 1$, 2 or 3, where $\nabla_x^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$ and ϖ denotes the pressure Let us multiply equation (41) by the k-component u'_k of the velocity at $P'(x'_1, x'_2, x'_3)$, so that

$$u'_k \frac{\partial u_i}{\partial t} + u'_k u_j \frac{\partial u_i}{\partial x_j} = \frac{-u'_k}{\rho}\frac{\partial \varpi}{\partial x_i} + \nu u'_k \nabla_x^2 u_i \tag{42}$$

Consider now the expression $u'_k u_j \frac{\partial u_i}{\partial x_j}$ It will be seen that, in virtue of the equation of continuity,

$$u'_k u_j \frac{\partial u_i}{\partial x_j} = \frac{\partial}{\partial x_j}(u_i u_j u'_k) \tag{43}$$

Now $\overline{u_i u_j u'_k}$ is a function of ξ_1, ξ_2 and ξ_3 and obviously

$$\frac{\partial}{\partial x_j}(\overline{u_i u_j u'_k}) = -\frac{\partial}{\partial \xi_j}(\overline{u_i u_j u'_k}) \tag{44}$$

Furthermore,

$$\frac{-u'_k}{\rho}\frac{\partial \varpi}{\partial x_i} = \frac{1}{\rho}\frac{\partial}{\partial \xi_i}(\varpi u'_k), \tag{45}$$

and since $\overline{\varpi u'_k}$ vanishes everywhere, $\overline{\frac{-u'_k}{\rho}\frac{\partial \varpi}{\partial x_i}}$ also vanishes Finally,

$$\nabla_x^2(\overline{u_i u'_k}) = \nabla^2(\overline{u_i u'_k}),$$

where

$$\nabla^2 = \frac{\partial^2}{\partial \xi_1^2} + \frac{\partial^2}{\partial \xi_2^2} + \frac{\partial^2}{\partial \xi_3^2}$$

Hence we obtain the equation

$$\overline{u'_k \frac{\partial u_i}{\partial t}} - \frac{\partial}{\partial \xi_j}(\overline{u_i u_j u'_k}) = \nu \nabla^2(\overline{u_i u'_k}) \tag{46}$$

In an analogous way we obtain

$$\overline{u_i \frac{\partial u_k'}{\partial t}} + \frac{\partial}{\partial \xi_j}(\overline{u_i u_j' u_k'}) = \nu \nabla^2 (\overline{u_i u_k'}) \tag{47}$$

We have shown in § 7 that

$$\overline{u_i u_j' u_k'} = -(\overline{u_i' u_j u_k}), \tag{48}$$

and substituting (48) in equation (47) we obtain

$$\overline{u_i \frac{\partial u_k'}{\partial t}} - \frac{\partial}{\partial \xi_j}(\overline{u_j u_k u_i'}) = \nu \nabla^2 (\overline{u_i u_k'}) \tag{49}$$

Adding the equations (46) and (49) and remembering that $\overline{u_i u_k'} = \overline{u^2} R_{ik}$ and $\overline{u_i u_j u_k'} = (\overline{u^2})^{\frac{3}{2}} T_{ijk}$, we obtain finally

$$\frac{\partial}{\partial t}(\overline{u^2} R_{ik}) - (\overline{u^2})^{\frac{3}{2}} \frac{\partial}{\partial \xi_j}(T_{ijk} + T_{kji}) = 2\nu \overline{u^2} \nabla^2 R_{ik} \tag{50}$$

In view of the form given for the tensors R_{ik} and T_{ijk} in equations (3) and (35) respectively, equation (50) may clearly be reduced to a differential equation connecting the functions f, g, k, q, h By using the equation of continuity in the form given by equations (7) and (39) we can eliminate g, k and q and obtain a partial differential equation connecting f and h The analysis can be carried out by choosing any component of R_{ik} Then equation (50) contains terms multiplied by $\xi_i \xi_k$ and terms which are functions of r alone Equating the terms containing $\xi_i \xi_k$ we obtain a relation for $\partial/\partial t\,(f-g)$ and equating the terms which are functions of r a relation for $\partial g/\partial t$ Eliminating $\partial g/\partial t$ the following equation for f results

$$\frac{\partial(f\overline{u^2})}{\partial t} + 2(\overline{u^2})^{\frac{3}{2}}\left(\frac{\partial h}{\partial r} + \frac{4}{r} h\right) = 2\nu \overline{u^2}\left(\frac{\partial^2 f}{\partial r^2} + \frac{4}{r}\frac{\partial f}{\partial r}\right) \tag{51}$$

We call (51) the fundamental equation for the propagation of the correlation function $f(r)$

9—*The decay of turbulence*

Before proceeding with the discussion of the solutions of equation (51) we shall show how Taylor's equation for the decay of energy and the equation for the decay of vorticity already given by the senior author (Kármán 1937*a*) can be obtained as deductions from equation (51)

$$\text{For } r = 0,\ f = 1 \text{ and } \frac{dh}{dr} + \frac{4}{r} h = 0 \quad (\text{cf } \S\, 6)$$

Hence we obtain from (51)

$$\frac{\partial \overline{u^4}}{\partial t} = 2\nu\overline{u^2}\left(\frac{\partial^2 f}{\partial r^2} + \frac{4}{r}\frac{\partial f}{\partial r}\right)_{r=0}, \tag{52}$$

or with

$$f = 1 + \tfrac{1}{2}f_0'' r^2 + \tfrac{1}{24}f_0^{\mathrm{iv}} r^4 + \quad , \tag{53}$$

$$\frac{d\overline{u^2}}{dt} = 10\nu f_0''\overline{u^2} \tag{54}$$

Now Taylor's definition of the length λ is given by the equation

$$\lambda^2 = -2/g_0'' \tag{54a}$$

We have shown (equation 10) that $f'' = \frac{1}{2}g_0''$ Hence

$$\frac{d\overline{u^2}}{dt} = -10\nu\frac{\overline{u^2}}{\lambda^2} \tag{55}$$

Or if, considering Taylor's problem of the decay of turbulence behind a honeycomb, we substitute $\frac{d}{dt} = \frac{1}{U}\frac{d}{dx}$,

$$\frac{1}{U}\frac{d}{dx}\overline{u^2} = -10\nu\frac{\overline{u^2}}{\lambda^2}, \tag{56}$$

which is Taylor's equation for the decrease of the mean kinetic energy of turbulence

Now we shall substitute the expansions given by equation (53) in equation (51) and equate the coefficients of the terms containing r^2 Concerning the triple correlation function $h(r)$ we notice that

$$h(r) = \tfrac{1}{6}h_0''' r^3 + \text{higher terms} \tag{57}$$

Hence we find

$$\frac{d}{dt}[\tfrac{1}{2}f_0''\overline{u^2}] + \tfrac{7}{3}h_0'''(\overline{u^2})^{\frac{3}{2}} = \tfrac{7}{3}\nu f_0^{\mathrm{iv}}\overline{u^2} \tag{58}$$

Obviously $f_0''\overline{u^2} = -\overline{u^2}/\lambda^2$ On the other hand, we find the mean square of the vorticity components using the formulae (19), (20) and (21)

$$\overline{\omega_1^2} = \overline{\omega_2^2} = \omega_3^2 = 5\overline{u^2}/\lambda$$

Hence equation (58) can be interpreted as the equation for the decay of vorticity Let us denote $\overline{\omega_1^2} + \overline{\omega_2^2} + \overline{\omega_3^2}$ by $\overline{\omega^2}$ then $\overline{\omega^2} = 15\overline{u^2}/\lambda^2$ and from (58) follows

$$\frac{d\overline{\omega^2}}{dt} - 70h_0'''(\overline{u^2})^{\frac{3}{2}} = -\frac{14}{3}\nu\overline{\omega^2}\lambda^2 f_0^{\mathrm{iv}}$$

$\lambda^2 f_0^{\text{iv}}$ is the reciprocal value of the square of a length, in order to obtain conformity with the energy equation (55) we put $\frac{1}{\lambda_\omega^2} = \frac{7}{15}\lambda^2 f_0^{\text{iv}}$ Then

$$\frac{d\overline{\omega^2}}{dt} - 70 h_0'''(\overline{u^2})^{\frac{3}{2}} = -10\nu\frac{\overline{\omega^2}}{\lambda_\omega^2} \tag{59}$$

From the equations of motion the senior author obtained the equation* (Kármán 1937*a*)

$$\frac{d\overline{\omega^2}}{dt} - 2\overline{\omega_i\,\omega_k\frac{\partial u_i}{\partial x_k}} = -10\nu\frac{\overline{\omega^2}}{\lambda_\omega^2} \tag{60}$$

Equations (59) and (60) are identical Because of the continuity equation

$$-2\,\overline{\omega_i\,\omega_k\frac{\partial u_i}{\partial x_k}} = \overline{\left(\frac{\partial^2(u_j u_i)}{\partial x_j \partial x_k} - \frac{\partial^2(u_j u_k)}{\partial x_j \partial x_i}\right)\left(\frac{\partial u_i}{\partial x_k} - \frac{\partial u_k}{\partial x_i}\right)}$$

Now the mean values, which constitute the expression on the right side, can be calculated by differentiation of the equation (35) For instance,

$$\overline{\frac{\partial^2(u_i u_j)}{\partial x_j \partial x_k}\frac{\partial u_i}{\partial x_k}} = \left[\frac{\partial^3\overline{(u_i u_j u_i')}}{\partial \xi_j \partial \xi_k \partial \xi_k}\right]_{\xi_i=\xi_j=\xi_k=0},$$

and so on

Carrying out the computations the identity of equations (59) and (60) can be shown without difficulty

It is seen that the equation for the decay of turbulence and for the decay of vorticity follow from the general equation for the spread of correlations by development of the correlation functions in powers of the distance *r* The expression for the decay of vorticity contains two terms one corresponds to the change of vorticity by deformation of the vortex tubes, the other corresponds to the action of viscosity

---

* As the senior author found the equation (60) he realized that the condition of isotropy alone does not lead to a further reduction of the equation However, he thought that in a random isotropic turbulence the expression containing the triple correlations should vanish because the extension and the contraction of portions of the vortex tubes should be equally probable in an ideal fluid As the authors carried out their first general analysis of correlations, it seemed that a general proof could be found for the vanishing of the triple correlations Correspondingly, the senior author stated in his papers on the subject (1937 *a*, *b*) that the triple correlations and the term in question in equation (60) vanish because of isotropy Closer analysis showed that this statement is erroneous To make the triple correlations vanish, some further physical assumption on the random character of turbulence is necessary

10—*Self-preserving correlation functions*

We write equation (51) in the form

$$\frac{\partial(f\overline{u^2})}{\partial t} = 2\nu\overline{u^2}\left(\frac{\partial}{\partial r}+\frac{4}{r}\right)\left[\frac{\partial f}{\partial r}-\frac{\sqrt{\overline{u^2}}}{\nu}h\right] \tag{61}$$

Obviously, the influence of the triple correlations depend on the relative magnitude of the quantities $\frac{\partial f}{\partial r}$ and $\frac{\sqrt{\overline{u^2}}}{\nu}h$ Let us consider first the case of "small Reynolds number" of the turbulence, i e let us assume that $\frac{\partial f}{\partial r} \gg \frac{\sqrt{\overline{u^2}}h}{\nu}$ Then neglecting the triple correlation function h, we write

$$\frac{\partial(f\overline{u^2})}{\partial t} = 2\nu\overline{u^2}\left(\frac{\partial^2 f}{\partial r^2}+\frac{4}{r}\frac{\partial f}{\partial r}\right), \tag{62}$$

or eliminating $\overline{u^2}$ by use of equation (54)

$$\frac{\partial f}{\partial t} = 2\nu\left(\frac{\partial^2 f}{\partial r^2}+\frac{4}{r}\frac{\partial f}{\partial r}-5f''_0 f\right) \tag{63}$$

We notice that in this case $f(r,t)$ is determined by the values of $f(r,t_0)$, i e if the correlation function f is given for $t=t_0$ it is given for any arbitrary time $t>t_0$

We shall consider a particular set of solutions of equation (63) We notice that this equation reduces to an ordinary differential equation if we suppose f is a function of $\chi = \frac{r}{\sqrt{(\nu t)}}$ only When this substitution is made, we obtain the equation

$$f''+\left(\frac{4}{\chi}+\frac{\chi}{4}\right)f'-5f''(0)f = 0, \tag{64}$$

where dashes denote differentiations with regard to χ, we shall speak of the correlation functions given by the solution of equation (64) as "self-preserving", since the form of these curves is the same at all instants although the actual length scale varies We shall now denote by α the arbitrary constant $-f''(0)$

First of all it may be pointed out that equation (64) is related to the confluent hypergeometric equation In fact, the solution we require (i e the one which satisfies $f=1$ and $f'=0$ when $\chi=0$) is equal to

$$f(\chi) = 2^{15/4}\chi^{-5/2}e^{-\chi^2/16}M_{(10\alpha-\frac{5}{4}),\frac{3}{4}}\left(\frac{\chi^2}{8}\right), \tag{65}$$

where $M_{k,m}(z)$ is the same solution of the confluent hypergeometric equation as that defined by Whittaker and Watson (1927, p 337, para 16 1) and denoted by this symbol

When $\alpha < \frac{1}{4}$ it will be seen, after some reduction, that (Whittaker and Watson 1927, p 352, example 1)

$$f(\chi) = \frac{\Gamma(5/2)}{\Gamma(10\alpha)\,\Gamma(\frac{5}{2}-10\alpha)} \int_0^1 \tau^{(10\alpha-1)}(1-\tau)^{(\frac{3}{2}-10\alpha)} e^{-\frac{\chi^2}{8}\tau}\,d\tau \tag{66}$$

When $\alpha = \frac{1}{4}$ we obtain the solution

$$f(\chi) = e^{-\chi^2/8} \tag{67}$$

When $\alpha > \frac{1}{4}$ the solution is given by the integral of the same integrand round a particular contour (Whittaker and Watson, 1927, p 257)

Let us now consider the decay of turbulence when the correlation function is one of these special types Returning to equation (54) we see that

$$\frac{\partial \overline{u^2}}{\partial t} = -10\frac{\overline{u^2}}{t}\alpha, \tag{68}$$

and integrating we find

$$\frac{1}{\sqrt{\overline{u^2}}} = \frac{1}{\sqrt{\overline{u_0^2}}}\left(\frac{t}{t_0}\right)^{5\alpha}, \tag{69}$$

where the condition $\overline{u^2} = \overline{u_0^2}$ when $t = t_0$ has been applied

Assuming, for the moment, that the turbulence produced by a particular grid or honeycomb is of the type which has a self-preserving correlation function and starting from some point outside the wind shadow, where the fluid passes at the time t_0, we write $t = t_0 + x/U$ (U denotes the velocity of the main flow and x is the distance measured downstream) Then we may write equation (69) in the form

$$\frac{1}{\sqrt{\overline{u^2}}} = \frac{1}{\sqrt{\overline{u_0^2}}}\left(1+\frac{x}{Ut_0}\right)^{5\alpha}, \tag{70}$$

where $\overline{u^2} = \overline{u_0^2}$ when $x = 0$

The quantity Ut_0 is a length which we can relate to the value λ_0 of the length λ (defined in equation (54a) above) at the origin from which x is measured For

$$\left(\frac{\partial^2 f}{\partial r^2}\right)_{r=0} = \frac{1}{\nu t}f''(0) = -\frac{\alpha}{\nu t},$$

so that, with $\left(\frac{\partial^2 f}{\partial r^2}\right)_{r=0} = -1/\lambda^2$

$$\lambda^2 = \frac{\nu t}{\alpha}, \tag{71}$$

and therefore $t_0 = \frac{\alpha\lambda_0^2}{\nu}$

Finally, from equation (70),

$$\frac{U}{\sqrt{\overline{u^2}}} = \frac{U}{\sqrt{\overline{u_0^2}}}\left(1 + \frac{x\nu}{\alpha\lambda_0^2 U}\right)^{5\alpha}, \tag{72}$$

where $\overline{u^2} = \overline{u_0^2}$, $\lambda = \lambda_0$ when $x = 0$

The results given by equations (71) and (72) would be formally equivalent with Taylor's results, when $\alpha = \frac{1}{5}$ However, the corresponding correlation function f which is given in fig 3 as calculated by numerical integration from the integral obtained by integrating (66) by parts is very

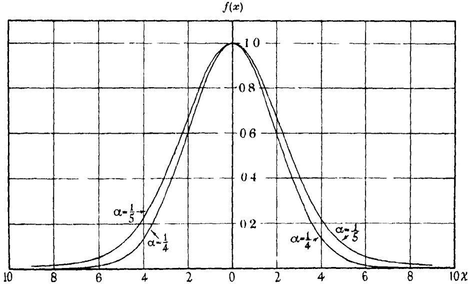


FIG 3

different from those measured by Simmons in such cases, in which Taylor's linear law of the decay of turbulence apparently holds Hence the coincidence between these equations and Taylor's result is rather formal We come back to this question in the last section of this paper The correlation function f corresponding to $\alpha = \frac{1}{4}$ (see equation (67)) is also included in fig 3

11—*Possible solution for large Reynolds numbers**

It was seen that if the triple correlations are neglected the double correlation function $f(r)$ is determined for all times $t > t_0$ provided $f(r)$ is known for $t = t_0$ In the general case $\partial f/\partial t$ depends also on the triple correlation function $h(r)$ Now a further equation can be obtained from the equations of motion for $\partial h/\partial t$, however, this new equation contains terms with the quadruple correlations and so on Hence, for an arbitrary value of the Reynolds number the problem is too complicated for analytical treatment

However, certain interesting results can be obtained for the limiting case of very large Reynolds numbers under the assumption that the correlation functions $f(r, t)$ and $h(r, t)$ are independent of viscosity with the exception of small values of the distance r Let us consider the problem of turbulence created by passing a uniform windstream through a grid or similar turbulence-producing device and let us assume geometrically similar arrangements Then, denoting the mean velocity of the windstream by U and a characteristic linear dimension of the device mentioned by M, it seems evident by dimensional considerations that f has the form

$$f\left(\frac{r}{\sqrt{(\nu t)}}, \frac{r}{M}, \frac{Ut}{M}\right)$$

If the correlation functions are—according to the above assumption—independent of viscosity, this means that $f(r, t)$ does not depend on the variable $\chi = \dfrac{r}{\sqrt{(\nu t)}}$ and is only a function of the variables r/M and Ut/M Let us now assume in addition that the correlation functions preserve their shape and only their scale is changing Obviously this assumption amounts to the statement that both $f(r, t)$ and $h(r, t)$ are functions of one variable $\psi = r/L$ only, where L is a function of M and Ut

It must be noticed that both assumptions—namely that the correlation functions are independent of χ and that they preserve their shape—shall be made only for large values of $\chi = \dfrac{r}{\sqrt{(\nu t)}}$

Then introducing $f(\psi)$ and $h(\psi)$ in equation (51), we obtain the following expression

$$-\frac{df}{d\psi}\psi\frac{\overline{u^2}}{L}\frac{dL}{dt}+f\frac{d\overline{u^2}}{dt}+2\frac{(\overline{u^2})^{\frac{3}{2}}}{L}\left(\frac{dh}{d\psi}+\frac{4h}{\psi}\right) = 2\nu\frac{\overline{u^2}}{L^2}\left(\frac{d^2f}{d\psi^2}+\frac{4}{\psi}\frac{df}{d\psi}\right) \qquad (73)$$

---

* This section was inserted by the senior author in Sept 1937, especially after reading G I Taylor's contribution (1937)

Let us consider the quantity $\frac{\sqrt{\overline{u^2}}L}{\nu}$ as the Reynolds number related to the problem, then for large values of this number it appears justified to neglect the term on the right side Furthermore, according to (55)

$$\frac{d\overline{u^2}}{dt} = -10\nu\frac{\overline{u^2}}{\lambda^2}$$

Hence, from (73) it follows that

$$-\frac{df}{d\psi}\psi\frac{\overline{u^2}}{L}\frac{dL}{dt} - f\frac{10\nu\overline{u^2}}{\lambda^2} + \left(\frac{dh}{d\psi} + \frac{4h}{\psi}\right)2\frac{(\overline{u^2})^{\frac{3}{2}}}{L} = 0 \tag{74}$$

Obviously equation (74) can be satisfied only when the coefficients which are functions of t without being functions of ψ, are proportional, i e their ratios are constants Hence

$$\frac{\sqrt{\overline{u^2}}\lambda^2}{L\nu} = A, \tag{75}$$

$$\frac{dL}{dt} = B\sqrt{\overline{u^2}}, \tag{76}$$

where A and B are numerical constants

We easily obtain from (75) and (76) a differential equation for $L(t)$ We substitute in (75) from (55) $\frac{\lambda^2}{\nu} = -10\overline{u^2}\Big/\frac{d}{dt}\overline{u^2}$ and eliminate $\overline{u^2}$ from (75) and (76) Thus we obtain

$$L\frac{d^2L}{dt^2} = -\frac{5}{AB}\left(\frac{dL}{dt}\right)^2 \tag{77}$$

The general solution of (77) is

$$L = L_0\left(1+\frac{t}{t_0}\right)^{\frac{5}{5+AB}}, \tag{78}$$

where L_0 and t_0 are arbitrary constants The origin of the time t being arbitrary, we may write

$$L = L_0\left(\frac{t}{t_0}\right)^{\frac{5}{5+AB}} \tag{79}$$

Introducing the expression (79) in (76) we obtain by differentiation

$$\sqrt{\overline{u^2}} = \frac{5}{5+AB}L_0\left(\frac{t}{t_0}\right)^{\frac{5}{5+AB}}\frac{1}{t}, \tag{80}$$

or denoting the value of $\sqrt{\overline{u^2}}$ at $t = t_0$ by $\sqrt{\overline{u_0^2}}$

$$\sqrt{\overline{u^2}} = \sqrt{\overline{u_0^2}}\left(\frac{t}{t_0}\right)^{\frac{-AB}{5+AB}} \tag{81}$$

Finally from (55)

$$\lambda^2 = \left(5+\frac{25}{AB}\right)\nu t \tag{82}$$

It is seen that equations (81) and (82) with $\frac{1}{\alpha} = 5+\frac{25}{AB}$ are identical with the corresponding equations (69) and (71) obtained in § 10 for the case of the small Reynolds number

There is one case which is not included in equation (77), namely when $L = \text{const}$ say $L = L_0$ Then from (75) with $\frac{\lambda^2}{\nu} = -5\sqrt{\overline{u^2}}\Big/\frac{d}{dt}\sqrt{\overline{u^2}}$ it follows that

$$\frac{1}{\sqrt{\overline{u^2}}} = -\frac{5}{AL_0}t \tag{83}$$

and

$$\lambda^2 = 5\nu t \tag{84}$$

Let us compare the results obtained with the researches of Taylor and Dryden Our last equations (83) and (84) are identical with Taylor's results, especially if we assume with Taylor that L_0 is proportional to M, which is in this case evident from dimensional consideration However, Taylor does not have the solutions (81) and (82) The reason is that Taylor found the equation (75) with a remarkable vision for the relations between the quantities involved, however, instead of (76) he assumed L proportional to M, i e a fixed ratio between the scale of turbulence and the linear size of the turbulence-producing device

In both cases—using Taylor's assumption or our broader relations (75) and (76)—the theory leads to the conclusion that the scale and distribution of turbulence in a windstream is independent of the speed of the windstream and so is the ratio $U/\sqrt{\overline{u^2}}$ measured at a certain distance x from the grid The difference is that according to our theory the scale of turbulence may increase downstream, while according to Taylor's assumption, it remains unchanged Instead of $L = \text{const} \times M$, we obtain in general $L = M \times$ function of Ut/M, and in the case of "self-preserving" turbulence the function involved is proportional to a certain power of Ut/M It appears that further experimental results will decide whether Taylor's assumptions are not too narrow and whether the equations (79) and (80) correspond more closely to the experimental facts

It is interesting that as far as the laws for the decay of turbulence and for the spread of correlation curves are concerned, our analysis of correlations leads essentially to the same results as the recently published theory of Dryden's (1937) which is based on entirely different conceptions It appears that the next step in the development of the theory should be to find the physical mechanism which is behind the mathematical relations (75) and (76), especially the mechanism which tends to increase the scale of turbulence without the action of the viscosity

REFERENCES

Dryden, H L 1937 *J Aero Sci* **4**, 273
Kármán, T de 1937 *a* *Proc Nat Acad Sci, Wash*, **23**, 98
— 1937 *b* *J Aero Sci* **4**, 131
Lamb 1932 "Hydrodynamics", 6th ed pp 208–9
Taylor, G I 1935 *Proc Roy Soc* A, **151**, 421–78
— 1937 *J Aero Sci* **4**, 311
Whittaker, E T and Watson 1927 "Modern Analysis" 4th ed p 337

The Effect of Temperature on the Photochemical Bleaching of Visual Purple Solutions

By H. J. A. Dartnall, C. F. Goodeve and R. J. Lythgoe

The Sir William Ramsay Laboratories of Inorganic and Physical Chemistry and the Department of Physiology and Biochemistry, University College, London

(*Communicated by Sir John Parsons, F.R.S.—Received* 5 *May* 1937)

1—Introduction

In a recent paper (Dartnall, Goodeve and Lythgoe 1936) a method has been described for the determination of the product of the extinction coefficient of visual purple, α, and the quantum efficiency of the photochemical bleaching, γ. It was shown that this product, $\alpha\gamma$, was a constant, independent of concentration over a wide range. It was also shown that the value of $\alpha\gamma$ was independent of light intensity and temperature over the narrow ranges studied. Kuhne (1879) stated that between 0° and 40° the light sensitivity of visual purple increased very little, and any changes which were present were confined to the first stage of "visual yellow" formation. Above 45° the rate of bleaching by light increased considerably. In very weak diffuse daylight a solution of visual purple bleached in 11 min. at 12° and in ½ min. at 50°. Amenomiya (1931), from a study of the retinae of living frogs, found temperature coefficients, per 10° rise, of 1·17 and 2·87 for the ranges 0–33° and 33–50° respectively. Hecht (1921) found that the velocity of bleaching with white light was independent of temperature over the range 5–36° for most of his experiments. In one series of experiments, however, a temperature coefficient of 1·15 was obtained.

In view of the theoretical importance of the temperature coefficient of photochemical reactions it was thought advisable to repeat and extend these observations using monochromatic light.

2—Technique

The arrangement of the apparatus was essentially the same as that described in the previous paper (Dartnall, Goodeve and Lythgoe 1936) with the exception of certain additions The solution of visual purple, which was contained in a cylindrical quartz cell having plane parallel faces, was bleached by a beam of light emerging from a monochromator A photoelectric cell was placed behind the solutions and connected to a galvanometer As the solution bleached, the light falling on the photoelectric cell increased and there were proportionate increases in the galvanometer deflexion It was thus possible to follow the course of the bleaching with time At the end of the experiment the cell containing the solution was removed from the holder and a calibrated thermopile substituted for the photoelectric cell The light intensity was then measured in absolute units (quanta per second) By subtracting an experimentally determined value for the fraction of the light reflected from the front surfaces of the cell, the light incident on the surface of the visual purple solution was obtained A small additional correction was made for the light reflected back into the solution from the back surface of the cell This correction increased as the solution bleached, and an average value was used to obtain the values in column (9) of Table I The wave-length of the monochromatic light was as before, 500–513 mμ with dominant wave-length 506 mμ

Temperature control

Additional arrangements were made to keep the temperature of the visual purple solution constant during an experiment and for a rapid change of the controlled temperature between experiments A stream of water from a thermostat was passed through the cell holder, a thermometer being placed in the exit stream The water normally returned directly to the thermostat The regulator in the thermostat was fitted with a calibrated side arm and tap for the rapid adjustment of the amount of mercury it contained and thus of the temperature to be maintained in the thermostat Two platinum wires were placed in the capillary of the regulator, one with its tip slightly above the other Contact of the mercury with the upper platinum wire operated an electromagnetic clip which directed the return flow from the cell holder through a copper spiral immersed in a stirred ice bath before returning to the thermostat—thus cooling the latter Breaking of the lower platinum-mercury contact turned on a heater in the thermostat In this way regulation was obtained at temperatures over the range 5–60°

The time taken for the contents of the cell to acquire a steady temperature and the relation between this temperature and that read on a thermometer in the water stream, was determined by means of a calibrated thermocouple immersed in the centre of the water-filled cell It was found, for example, that with the cell initially at 0° and a flow-water temperature of 39 1°, a steady temperature of 38 0° was set up inside the cell in 10 min In all experiments a suitable interval was allowed to obtain the steady state, care being taken to avoid too much thermal decomposition at the higher temperatures The corrections being rather variable and, as will be seen below, an accurate knowledge of the temperature being unnecessary, the temperatures recorded in this paper are those of the flow-water

Method of extraction

The visual purple was extracted from the retinae of frogs The method of extraction is summarized here, the full details being given in a paper by Lythgoe (1937)

The rods containing the visual purple were removed from the retina and dried They were then washed with petrol ether (B P 40–60°) to remove ether-soluble impurities and treated with a *p*H 4 6 buffer solution to precipitate some of the proteins Finally the visual purple was liberated from the rods by adding a solution of 1 % digitonin in water After centrifuging, the supernatant fluid was buffered by the addition of an equal volume of M/20 standard Clark and Lubs' buffer solution (Clark 1928) The final *p*H values of the solutions were determined or calculated All solutions prepared by this method were free from opalescence and remained so during the bleaching They did not form any precipitate even after they had been kept for several weeks The much higher purity of the visual purple solutions used for the present experiments can be appreciated by comparing the densities after bleaching (Table I, p 226) with those in the previous paper

Rana esculenta were used for all of the experiments except Nos 61 and 62, in these, ten *R temporaria* were used with two *R esculenta* Usually twelve frogs were used to prepare 2 c c of the final solution The experiments were carried out in the summer and early autumn of 1936

3—THEORETICAL

The following symbols and definitions are used in this paper

I, intensity of light, expressed as number of quanta per second, incident on the solution,

I_t, intensity of light transmitted by the solution at a time t,

I_f, intensity of light transmitted by the solution after complete bleaching,

t, time from the initial exposure of the solution to the light,

l, internal length of the optical cell,

A, exposed area of the solution,

V, volume of the solution,

n, number of centres of absorption (chromophoric groups) in the volume V of solution,

c, concentration, expressed as the number of chromophoric groups divided by the volume of solution, thus $c = n/V$,

γ, quantum efficiency, a ratio defined as

$$\frac{\text{number of chromophoric groupings destroyed}}{\text{number of quanta absorbed by visual purple}},$$

α, extinction coefficient per chromophoric group expressed in square centimetres and defined by the equation

$$\log_e \frac{I}{I_t} = \alpha c l$$

In a previous paper (Dartnall, Goodeve and Lythgoe 1936) we have shown that the theoretical equation

$$\log_e \frac{I_t}{I_f - I_t} = \phi \, \frac{\alpha \gamma I}{A} \, t + \text{constant} \tag{1}$$

accurately describes the bleaching kinetics of visual purple solutions containing stable light absorbing impurities Experimental values of

$$\log_{10} \frac{I_t}{I_f - I_t},$$

when plotted against time gave straight lines of slope $\phi \, \dfrac{\alpha \gamma I}{2\,3A}$

The quantity ϕ has been shown to have the value,

$$\phi = \frac{I_f}{I_f - I_t} \, \frac{I - I_t}{I} \, \frac{\log_e I_f/I_t}{\log_e I/I_t}, \tag{2}$$

but, although a variable, the change in its value for any one experiment is negligible The rate of bleaching of a solution containing foreign absorbing matter would be less than that for a pure solution The slope of the straight line would be less, but the value of ϕ gives an exact measure of this reduction Correct values for $\alpha\gamma$ in absolute units could, therefore, be obtained For pure solutions of visual purple, $I_f = I$, and $\phi = 1$

In the derivation of equation (1) it was assumed that the products of the bleaching do not absorb light appreciably Such products would not affect the validity of the equation if they decomposed to a colourless substance at a rate rapid compared with the rate of photodecomposition of visual purple This, however, is not always so and it will, therefore, be necessary to consider cases where the light absorption of the products cannot be neglected

The case where stable light absorbing substances are produced

It will be assumed that the products of the bleaching are formed stoichiometrically from the decomposition of visual purple and that once formed they are stable Assuming that Beers's and Lambert's laws of light absorption hold for all components, the optical density of the solution at any time, t, is

$$\log_e \frac{I}{I_t} = \alpha c l + z\alpha'(c_0 - c)\, l + d_i, \tag{3}$$

where $z\alpha'$ has been written for $a\alpha_a + b\alpha_b +$, α_a, α_b, , being the extinction coefficients of the products of the bleaching and a, b their stoichiometric relation to visual purple, d_i, the density of the impurities (stable) present and c_0, the concentration when $t = 0$

Differentiating (3) with respect to time and rearranging,

$$-\frac{dI_t}{I_t\, dt} = (\alpha l - z\alpha' l)\frac{dc}{dt} \tag{4}$$

From the definition of quantum efficiency given above, the rate of decrease in the number of visual purple chromophoric groups, $-\frac{dn}{dt}$, is equal to the product of the quantum efficiency and the intensity of the *light absorbed by the visual purple* This latter quantity is related to the total light absorbed $(I - I_t)$ by the ratio of $\alpha c l$ to the total density We have, therefore,

$$-\frac{dn}{dt} = -V\frac{dc}{dt} = \gamma\, \frac{\alpha c l}{\alpha c l + z\alpha'(c_0 - c)\, l + d_i}\, (I - I_t) \tag{5}$$

By eliminating dc/dt between (4) and (5) and rearranging,

$$\frac{dI_t}{I_t\, dt} = \frac{\alpha c l - z\alpha' c l}{\alpha c l + z\alpha'(c_0 - c)\, l + d_i}\, \frac{I - I_t}{I}\, \frac{\alpha \gamma I}{A} \tag{6}$$

The density of the completely bleached solution is

$$\log_e \frac{I}{I_t} = z\alpha' c_0 + d_i \tag{7}$$

Subtracting equation (7) from equation (3)

$$\log_e \frac{I_f}{I_t} = \alpha cl - x\alpha' cl \tag{8}$$

Whence, substituting equations (3) and (8) in (6)

$$\frac{dI_t}{I_t dt} = \frac{\log_e I_f/I_t}{\log_e I/I_t} \frac{I - I_t}{I} \frac{\alpha\gamma I}{A} \tag{9}$$

Multiplying both sides of this equation by $\frac{I_f}{I_f - I_t}$, and comparing with equation (2),

$$\frac{I_f}{I_f - I_t} \frac{dI_t}{I_t dt} = \phi \frac{\alpha\gamma I}{A} \tag{10}$$

As shown below, ϕ may be considered constant Therefore, integrating (10),

$$\log_e \frac{I_t}{I_f - I_t} = \phi \frac{\alpha\gamma I}{A} t + \text{constant} \tag{11}$$

It is seen that this equation is identical with equation (1) The plot of $\log \frac{I_t}{I_f - I_t}$ against t should therefore give a straight line as in the cases considered previously

The correcting function ϕ

From equation (2) it is seen that ϕ depends on two constants, I and I_f, and on a variable, I_t These may be expressed in any convenient units such as galvanometer deflexions (when linear with I) I_t may have values anywhere between I and very nearly zero, and may be greater than or less than I_f Values of I_t greater than I_f correspond to the case where the photochemical change causes an increase in the density due to the formation of a stable product which absorbs light more strongly than the original substance This occurs with visual purple in the violet end of the spectrum

At the end of an experiment, that is when no further change takes place, I_t is equal to I_f and ϕ is given by

$$\phi = \frac{I - I_f}{I} \frac{1}{\log_e I/I_f} \tag{12}$$

In all of the experiments described below the value of ϕ remained remarkably constant, depending almost exclusively upon the values of I and I_f

according to equation (12). This can be shown always to be the case by taking any arbitrary set of values of I, I_f, and I_t and calculating ϕ.

The maximum variations in ϕ over the whole range of a bleaching curve with an initial optical density of 0·5, for various values of I_f/I are as follows

I_f/I	1·0	0·9	0·8	0·5
Final density	0·0	0·046	0·097	0·301
Value of $\phi_{t=0}$	1·0	0·957	0·91	0·74
Value of $\phi_{t=\infty}$	1·0	0·95	0·897	0·721
Variation	0	0·7 %	1·4 %	2·7 %

For most of the experiments described below, I_f/I was between 0·9 and 1·0.

The correction for the thermal decomposition of visual purple

Visual purple decomposes in the dark, and at temperatures above 25° corrections for this thermal reaction were necessary. A number of experiments were carried out and it was found that at a pH of 9·2 the velocity of decomposition followed the equation of a first order reaction

$$-\frac{dc}{dt} = k_1 c \text{ or } \log_{10} c = -\frac{k_1 t}{2·3} + \log_{10} c_0 \qquad (13)$$

The velocity constant k_1, expressed in sec $^{-1}$, increased with the absolute temperature, T, approximately according to the equation,

$$\log_{10} k_1 = -\frac{8{,}670}{T} + 23·9 \qquad (14)$$

The thermal reaction was found to give the same products as the photochemical reaction. The reactions presumably proceed independently and, therefore, the total reaction rate on illumination is given by equation (5) plus an extra term, "$k_1 Vc$". By proceeding as above one arrives at the following equation in place of equation (11)

$$\log_e \frac{I_t}{I_f - I_t} = \phi \frac{\alpha\gamma I}{A} t + k_1 \int_{t=0}^{t} \frac{I_f}{I_f - I_t} \log_e \frac{I_f}{I_t}\, dt + \text{constant} \qquad (15)$$

For solutions of not too high a concentration of visual purple the integral term becomes simply $k_1 t$. In such cases the correction for the thermal decomposition can be made by subtracting k_1 from the slope of the $\log_e \frac{I_t}{I_f - I_t}$ time line. For high concentrations straight lines should not be obtained if the thermal decomposition is pronounced.

In most of the experiments indicated in Table I this correction could be justifiably applied as the densities were not high and the correction itself was not large For a few experiments at pH's other than 9 2 special determinations of the correction were made

The coloured products of photodecomposition of visual purple

It has been generally accepted that visual purple bleaches to a substance called visual yellow It has been found (Lythgoe 1937) by a study of absorption spectra that at least two coloured products may arise, one of which has been called "transient orange" from its colour and instability at temperatures above 20° and a second which has been called "indicator yellow" from the fact that it changes colour with pH Indicator yellow has a deep yellow colour in acid, and a pale lemon colour in alkaline solutions It is not very stable in acid but is stable up to 50° in weak alkaline solutions Most of the experiments referred to below were carried out at a pH of about 9 2 where the absorption by indicator yellow of light of $\lambda = 506\,\mathrm{m}\mu$ is negligible

On bleaching rapidly a solution of visual purple at a temperature of about 10°, the colour of transient orange appears, and passes ultimately to that of indicator yellow It is possible (Lythgoe 1937) that transient orange is intermediate between visual purple and indicator yellow, being formed from the former by the action of light and decomposing into the latter thermally No evidence of a photochemical decomposition of transient orange has been obtained

RESULTS

The experimental values of $\log_{10} \frac{I_t}{I_f - I_t}$ for twenty-nine experiments are plotted against the time in the figure The time scales of the curves have been successively displaced merely for convenience, the start of the illumination in each case being indicated by the short vertical line Experiments carried out with the same preparation of visual purple are grouped together All experiments carried out are shown, with the exceptions of No 55 which was used for special purposes and of No 75 which failed for mechanical reasons In most cases the first observation of the light intensity was made one half-minute after the beginning of the illumination and thereafter at minute intervals Observations were carried out until the solution was completely bleached but those very near the end are not

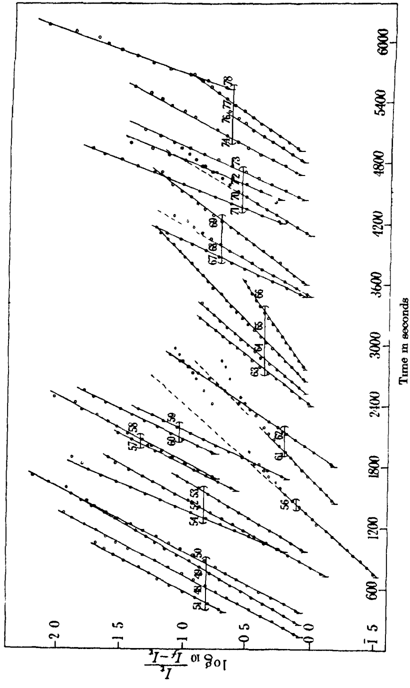

$\log_{10} \frac{I_t}{I_f - I_t}$
2·0
1·5
1·0
0·5
0·0
$\bar{1}$·5
600
1200
1800
2400
3000
3600
4200
4800
5400
6000
Time in seconds
48
49
50
51
52
53
54
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
76
77
78

FIG 1

plotted in fig 1 as the corresponding values of the function $\log \frac{I_t}{I_f - I_t}$ have an error grossly disproportionate to that of the light intensities themselves

The important data for the experiments are given in Table I, the first section of which refers to those experiments carried out at the alkaline *p*H of about 9 2, and the second to those carried out in approximately neutral solutions The experimental data are arranged in order of increasing temperatures, the values of which are given in column (10) The initial densities $\left(\log_{10} \frac{I}{I_{t=0}}\right)$, column (3), correspond to the lower terminus of each curve The final densities $\left(\log_{10} \frac{I}{I_f}\right)$, column (4), indicate the light absorption due to impurities and/or the products of bleaching

It will be seen from the figure that the results of all experiments, except Nos 56, 61, 68 and 70, lie closely on straight lines This may be taken to indicate that the above theoretical treatment adequately describes the bleaching process in the twenty-five experiments, and that in each case the product of the extinction coefficient and the quantum efficiency, $\alpha\gamma$, may be calculated from the slope of the straight line The values of the slopes (with the factor for conversion to natural logarithms) are given in column (5) Where there was a slight departure from the straight line more weight was given to the slope in the earlier part of the experiment The first order velocity constants for the thermal reaction are given in column (6) On subtracting these from the values in column (5) (see equation (15) *et seq*) the photochemical contribution to the slope is obtained as set out in column (7) The mean values of ϕ, equation (2), and the light intensities in quanta per sec are given in columns (8) and (9) respectively The cross-sectional area of the cells used for all experiments was 0 913 cm ^{2} The values of $\alpha\gamma$ calculated from these results are given in column (11)

The experiments, Nos 56, 61, 68 and 70 for which straight lines were not obtained, were all carried out between 10° and 15°, whereas it will be noted that straight lines were obtained at 5° and at and above 20° As described earlier, one of the products, "transient orange", has been found to be unstable at temperatures above 10° Above 20° its rate of thermal decomposition is very fast compared with the rate of photodecomposition of visual purple and, therefore, its concentration never becomes great enough to have an appreciable effect on the light absorption At 5° it is comparatively stable and its effect is accounted for by the value of ϕ according

Time in seconds

TABLE I

(1) Exp no	(2) pH	(3) Initial density	(4) Final density	(5) Slope × 10^4 (sec $^{-1}$)	(6) $k \times 10^4$ (sec $^{-1}$)	(7) Corrected slope × 10^4 (sec $^{-1}$)	(8) ϕ mean	(9) Intensity × 10^{-13} quanta/sec	(10) Temp °C	(11) $\alpha\gamma \times 10^{17}$ (cm ^{2})
65	9·2	0·308	0·031	18·9	0·0	18·9	0·967	2·29	5·0	7·8
69	9·2	0·314	0·042	24·6	0·0	24·6	0·955	3·25	5·0	7·2
68	9·2	0·315	0·021	33·7	0·0	33·7	0·977	3·67	10·0	(8·6)
61	9·2	0·308	0·022	19·4	0·0	19·4	0·977	3·34	10·7	(5·4)*
56	9·2	0·808	0·115	20·9	0·0	20·9	0·885	3·66	14·6	(5·9)
70	9·2	0·309	0·021	30·1	0·0	30·1	0·977	4·19	15·0	(6·7)
67	9·2	0·288	0·006	41·1	0·0	41·1	0·993	3·91	20·2	9·7
48	9·2	0·276	0·004	35·6	0·0	35·6	0·996	3·54	21·7	9·2
49	9·2	0·283	0·022	32·1	0·1	32·0	0·976	3·44	24·4	8·7
63	9·2	0·310	0·017	23·5	0·1	23·4	0·982	2·33	27·4	9·3
50	9·2	0·308	0·043	35·5	0·2	35·3	0·954	3·35	29·2	10·1
52	9·2	0·390	0·019	34·3	0·2	34·1	0·980	3·53	30·2	9·0
51	9·2	0·113	0·028	35·8	0·4	35·4	0·969	3·29	33·4	10·1†
73	9·2	0·288	0·022	41·1	0·6	40·5	0·977	4·00	35·2	9·3
62	9·2	0·438	0·035	27·1	0·6	26·5	0·964	3·01	35·3	8·4*†
57	9·2	0·126	0·027	37·8	1·6	36·2	0·970	3·7	40·0	9·2
58	9·2	0·115	0·033	35·1	1·6	33·5	0·966	3·67	40·0	8·7
53	9·2	0·321	0·039	31·6	1·8	29·8	0·958	3·37	40·6	8·5
59	9·2	0·249	0·036	39·5	4·4	35·1	0·961	3·82	45·0	8·7
60	9·2	0·109	0·041	38·6	4·4	34·2	0·955	3·77	45·0	8·7†‡
54	9·2	0·309	0·092	44·5	7·1	37·4	0·905	3·34	47·5	11·3‡
72	9·2	0·212	0·008	45·3	7·1	38·2	0·991	4·04	47·5	8·7
71	9·2	0·212	0·010	47·5	11·5	36·0	0·989	3·93	50·0	8·5
76	7·8	0·333	0·069	28·0	0·0	28·0	0·927	3·31	25·0	8·3
74	7·8	0·308	0·058	34·9	0·2	34·7	0·938	3·67	35·0	9·2
66	6·8	0·312	0·043	12·8	0·0	12·8	0·954	2·17	5·0	5·7
77	6·8	0·340	0·090	27·1	0·0	27·1	0·907	3·12	25·0	8·7
64	6·8	0·303	0·029	22·1	0·1	22·0	0·969	2·36	27·4	8·8
78	6·8	0·142	0·050	54·4	23 approx	31·4	0·945	3·2	60·0	9·4

* Ten *Rana temporaria* with two *R. esculenta*

† Solution considerably bleached prior to experiment

‡ I_f obtained by short extrapolation

to the above theoretical treatment At temperatures between 10° and 15° where its rate of decomposition is comparable with that of visual purple, straight lines are not obtained and the above method of analysis breaks down Very approximate values for $\alpha\gamma$, obtained from values of I_f which were calculated assuming that the transient orange was stable, are shown in brackets in column (11) It is to be noticed that the value of $\log_{10}\frac{I_t}{I_f - I_t}$ for each of these four experiments falls below the straight line in the latter part of the experiment This is to be expected from the fact that the presence of transient orange has slowed down the rate of attainment of the final intensity of the transmitted light The range of temperature over which straight line relations are not expected depends on the light intensity At higher light intensities the bleaching of visual purple may be comparable with the rate of decomposition of the transient orange at 20° or even at 25° In very weak light, on the other hand, transient orange may not be observed at the low temperatures The discrepancies in the literature with regard to the colour of the products from visual purple may be bound up with these facts

It has been found by Kuhne (1879) that a frozen retina ($-13°$ C) bleaches much more slowly than an unfrozen retina whose temperature was just below 0° C Garten (1906) confirmed these observations We also have found that frozen solutions of visual purple bleach very slowly but it was not possible to carry out quantitative experiments of this type

Discussion

The straight lines shown in the figure provide extensive confirmation of the statement made in the previous paper that the value of $\alpha\gamma$ is independent of concentration over a wide range α is almost certainly a constant at these very low concentrations, and, therefore, one may conclude that the quantum efficiency is independent of concentration

From the values in column (11) of Table I it is seen that $\alpha\gamma$ is independent both of temperature, except possibly at low temperatures, and of *p*H, over the range 6 8–9 2 Separate measurements have shown that α is effectively constant (decreasing by about 2%) between 18° and 33° and is independent of *p*H over a wide range One may, therefore, conclude that the quantum efficiency is also independent of temperature and *p*H The lower values obtained at 5° may be taken to indicate a slight fall of the quantum efficiency, which fact may be connected with the probable low value of $\alpha\gamma$ in a frozen solution

The mean value of $\alpha\gamma$ for $\lambda = 506\,m\mu$ over the range 20–60° is found to be 9×10^{-17} cm ^{2} molecules per quantum In taking this mean, the values in brackets, those at 5° and that for Exp 54, have been neglected The high value for Exp 54 is inexplicable—possibly a catalyst was present which caused a faster thermal reaction than that calculated from equation (14)

The accuracy of the above absolute value for $\alpha\gamma$ depends on the determination of the light intensity by means of a thermopile Unfortunately this measurement was subject to a large but nearly constant correction due to the presence of infra-red radiation reflected by the optical parts of the monochromator The correction has been determined and applied to the results in Table I, but they may have a constant absolute error of $\pm$ 20% The results given in the previous paper were not corrected, but when this is done they are brought into agreement with those given here

CONCLUSIONS

A photochemical reaction may be divided into two stages The first, the ' primary process", consists in the absorption of a quantum of light resulting in the production of an activated molecule, which may, in certain cases, be a new molecular species Temperature is invariably without effect on this part of the reaction and one and only one molecule or chromophoric group is altered by the absorption of a light quantum

"Secondary processes" are of a number of types In some photochemical reactions, the products of the primary process catalyse further decomposition of the photo-reactant The mechanism is commonly of the type known as a "chain reaction" Such secondary processes almost invariably are markedly influenced by temperature and commonly by concentration They lead to quantum efficiencies for the whole process greater than unity and varying over a wide range The constancy of γ for the photo-decomposition of visual purple provides a strong argument against the existence in this reaction of secondary processes of this type One may therefore conclude that the quantum efficiency for the bleaching of visual purple is not greater than unity This leads to the conclusion that the extinction coefficient α is equal to or greater than 9×10^{-17} cm ^{2}

In other reactions, activated molecules may be deactivated by collisions with solvent molecules before the photochemical change has time to become complete The number of molecules bleached by the light becomes only a fraction of the number of quanta absorbed, i e γ becomes less than unity The fraction of the primary process which is followed by deactivation is generally found to be independent of concentration at low concentrations and more or less independent of temperature

The completion of the photochemical change acts in competition with the deactivation This completion may involve a rearrangement of the atoms in the activated molecule, a reaction with the solvent water or a reaction with another part of the molecule, etc The nature of this part of the photochemical process cannot be profitably discussed on existing evidence It can be said, however, that, if the completion takes place very quickly after the primary process, temperature would have little influence If, on the other hand, it is slow due to the existence of an energy of activation, temperature would have a marked effect

It is difficult to estimate the relative probability of deactivation and of completion except that one could say that a high deactivation fraction would not be easily reconciled with the observed constancy of the photochemical bleaching of visual purple when it is subjected to variations of temperature, of pH and of buffer salt concentration On the other hand, considerable deactivation must take place in frozen solutions and possibly some at 5° It would seem simpler to conclude that deactivation is absent above 20° and sets in at 5° becoming strong in the frozen solution, rather than that an otherwise constant deactivation fraction suddenly increases when the temperature is lowered below 5° Absence of deactivation would mean that the absorption of a quantum of light always resulted in a permanent photochemical change, i e that γ would be equal to unity The extinction coefficient α for $\lambda = 506\,\mathrm{m}\mu$ would have the value, 9×10^{-17} cm ^{2} and ϵ, the decadic molar extinction coefficient, the value, $2{\cdot}3 \times 10^4$ Such a value is a normal one for a chromophoric group which exhibits *continuous* absorption of the same type as visual purple In the previous paper (Dartnall, Goodeve and Lythgoe 1936) this was put forward as an argument in favour of a low deactivation fraction (a high fraction would lead to a low value of γ and an abnormally high value of α) Further searches of the literature have served to strengthen this argument as no substances exhibiting this type of absorption have been found with a maximum of extinction more than twice that of visual purple

The above arguments, based on the experiments described in this paper, lead to the conclusion that the quantum efficiency for the bleaching of visual purple is equal to or not much less than unity

SUMMARY

The theoretical considerations developed previously for the photochemical bleaching of a light absorbing substance have been extended to include the case where the products absorb light Twenty-five out of twenty nine experiments with visual purple were found to be in accordance with these and the previous considerations In the remaining four, the decomposition of a coloured product, "transient orange", proceeded at a rate comparable with the rate of bleaching of visual purple The quantum efficiency of the bleaching process was found to be independent of temperature over the range, 20–60°, and of *p*H from 6 8 to 9 2 It is concluded that the quantum efficiency is equal to or not much less than unity, and that the molar extinction coefficient ($\lambda = 506\,\mathrm{m}\mu$) is $2\ 3 \times 10^4$

The authors are indebted to Mr R C Amsden, Mr J P Quilliam and Dr E Schneider for assistance with the experiments and to the Rockefeller Foundation for a generous grant in aid of this research

REFERENCES

Amenomiya, J 1931 *Acta Soc ophthal jap* **35**, 22, see *Zbl ges Ophthal* **25**, 250
Clark, W M 1928 ' The Determination of Hydrogen Ions " London
Dartnall, H J A , Goodeve, C F and Lythgoe, R J 1936 *Proc Roy Soc* **A, 156**, 158
Garten, S 1906 *v Graefes Arch Ophthal* **63**, 112
Hecht, S 1921 *J Gen Physiol* **3**, 285
Kuhne, W 1879 In Hermann 'Handbuch Physiol ' **3**, Part I, 284
Lythgoe, R J 1937 *J Physiol* **89**, 331–58

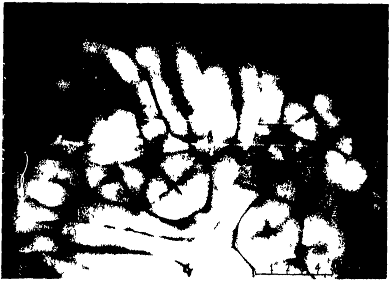


Fig 2—10 mm depth with temperature difference 117° C

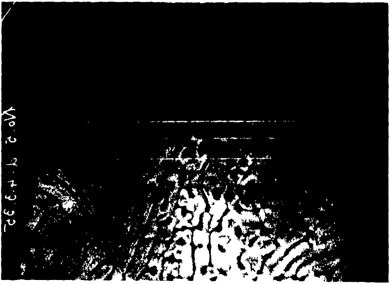


Fig 3—6 mm depth, with temperature difference 112° C

Fig 4—7 mm depth, with temperature difference 130° C

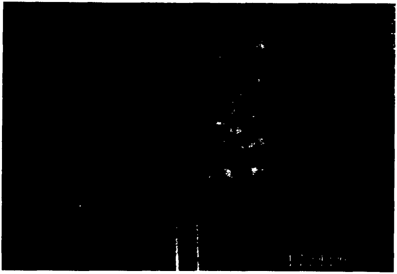


Fig 5—8 mm depth, with temperature difference 90° C, and shear 1·7 cm/sec

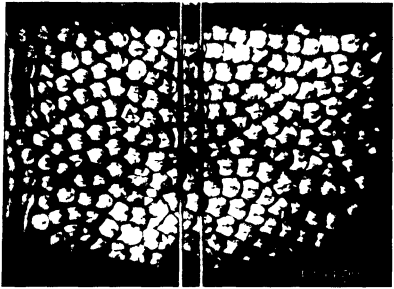


Fig 6—7 mm depth with temperature difference 103 C and shear 2 3 cm/sec

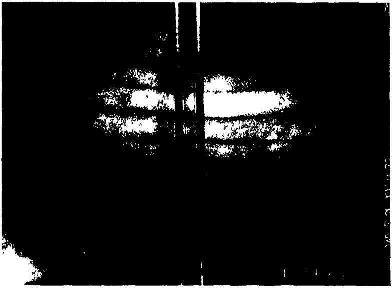


Fig 7—16 mm depth, with temperature difference 29 C, and shear 10 cm/sec

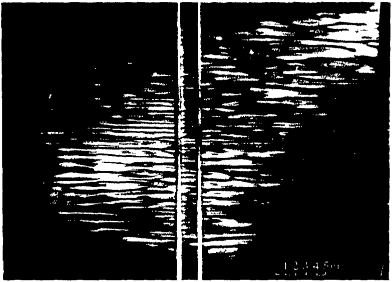


Fig. 8—6 mm depth with temperature difference 91° C and shear 2·3 cm/sec

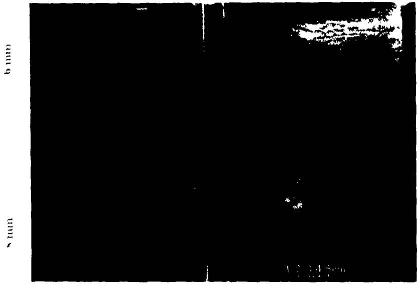


Fig. 9 Depth varying continuously from 8 mm at bottom to 6 mm at top of picture, with temperature difference 42° C, after shear was stopped

Instability of Fluids Heated from Below*

By Krishna Chandra, M Sc

(*Communicated by Sir Gilbert T Walker, F R S —Received* 20 *May* 1936
—*Revised* 12 *November* 1937)

[Plates 3–6]

Introduction

A considerable amount of research has been carried out in this country and elsewhere on the forms of motion in layers of fluid made unstable by heating from below of which a summary has been given by Walker (1935), and a theoretical treatment by the late Lord Rayleigh (1916) has given an indication of the magnitudes of the temperature differences which can subsist in a layer of fluid without motion Jeffreys (1928) has also discussed the theoretical aspects of the question, and modified the criterion given by Rayleigh for a variety of boundary conditions

The present research was undertaken at the suggestion of Professor D Brunt with a view to testing the criterion of stability given by Jeffreys, according to which the difference of temperature between the top and bottom of a layer necessary to produce motion in the layer is approximately inversely proportional to the cube of the depth, and to studying the variation in the form of the motion with varying depth of the layer of fluid and with varying differences of temperature between top and bottom of the layer

In Part I is discussed the case where the fluid as a whole has no initial motion, while Part II refers to the case where the fluid is sheared by the motion of its upper boundary The present research is a continuation of the work carried out in the Meteorological Department at the Imperial College of Science and Technology, London, by Sir Gilbert Walker and a number of his pupils

Part I

1—*Experimental apparatus*

The experimental chamber—The base of the chamber consisted of a flat stainless steel plate, of thickness $\frac{1}{4}$ in and area 12×12 in The plate rested on a hollow asbestos-lined box, 8 in in height, having the same area as the plate and provided with three levelling screws The metal plate could be heated from below by passing a current through a number of coils of thin wire arranged at the bottom of the box

---

* Part I is an abstract of a thesis for which the degree of M Sc was awarded by the University of London

The top of the working chamber was a thick sheet of plate glass 15 in square, fitted in a wooden frame in such a way that the glass formed the bottom of a vessel which could be filled with cold water By this means the temperature of the top of the vessel could be kept approximately constant Sets of brass pillars were used to support the glass sheet at any desired height above the metal plate The sides of the chamber consisted of a number of felt borders, introduced between the metal plate and the glass top A two-way hand-pump was used for producing smoke from a cigarette and introducing it into the chamber as an indicator of the motion

Temperature measurement inside the chamber—Temperature measurements were made by means of platinum resistance thermometers, the resistance elements being of diameter 0 004 in Two of these elements were fixed as close as possible to the top and bottom plates, while another was fixed half-way between the plates

2—*Experimental test of Jeffreys' formula*

Lord Rayleigh (1916), in a mathematical discussion of the breakdown of stability in layers of fluid heated from below, found that it was possible for a layer of fluid to remain in equilibrium, with denser fluid at the top than at the bottom, so long as the condition

$$\frac{\rho_1-\rho_0}{\rho_0}<\frac{27\pi^4\kappa\nu}{4gh^3}$$

is fulfilled Here ρ_1 is the density of the fluid at the top, ρ_0 the density at the bottom, κ the coefficient of molecular diffusivity of heat, ν the kinematic coefficient of viscosity, g the acceleration of gravity, and h the depth of the layer For a gas this condition can be expressed in terms of temperature as follows

$$\frac{T_0-T_1}{T}<\frac{27\pi^4\kappa\nu}{4gh^3} \quad \text{or} \quad \frac{658\kappa\nu}{gh^3},$$

where T_0 and T_1 are the temperatures at the bottom and top of the chamber, respectively, and T the mean temperature

Jeffreys (1928) showed that for the boundary conditions of the present series of experiments the above inequality should be replaced by

$$\frac{T_0 - T_1}{T} < \frac{1709\kappa\nu}{gh^3}, \quad \text{or} \quad \frac{T_0 - T_1}{\kappa\nu T} < \frac{1709}{gh^3}$$

It was readily shown experimentally that the layer of air could remain in equilibrium with higher temperature at the bottom than at the top. When the bottom plate was heated, no motion was observed in the chamber until the difference of temperature between bottom and top exceeded a limiting value whose magnitude depended on the depth of the layer. Thus, when the depth was 10 mm., no motion could be observed until the difference of temperature exceeded about 11.4° C.

In practice it was found easier to estimate the limiting value by first heating the bottom plate until the difference of temperature was sufficient to produce motion, then allowing the metal plate to cool slowly. The critical stage at which all motion had just ceased could then be observed, and the corresponding difference of temperature measured. Smoke was introduced into the chamber at intervals, and the patterns of the motion closely watched. As the difference of temperature gradually diminished, the motion in the chamber slowed down, and eventually a stage was reached at which it stopped completely. At this stage, as well as with still lower differences of temperature, no patterns of flow could be seen when fresh smoke was introduced into the chamber. As the critical stage was approached, temperature readings were taken in quick succession, and the temperature difference at the critical stage could be determined with reasonable accuracy.

The critical temperature differences for various depths are given in Table I, in the column headed ΔT. The next column gives $\Delta T/\kappa\nu T$. The

TABLE I—DIFFERENCES OF TEMPERATURE BETWEEN TOP AND BOTTOM OF A LAYER OF AIR AT CRITICAL STAGE WHEN MOTION IS JUST POSSIBLE

Depth mm	Temp bottom ° C c	Temp top ° C a	Temp centre ° C b	ΔT ° C $c-a$	$\Delta T/\kappa\nu T$	$1709/gh^3$
16	30.2	26.7	29.5	3.5	0.32	0.41
12	36.7	27.6	32.3	9.1	0.81	0.99
10	38.3	26.8	32.1	11.4	1.01	1.71
8	33.8	23.7	28.5	10.1	0.91	3.33
7.5	33.6	25.8	28.6	7.8	0.70	4.03
7	32.3	24.8	28.5	7.5	0.68	5.00
6	34.3	27.3	30.9	7.0	0.62	7.90
4	32.3	27.0	30.2	5.3	0.48	26.70

corresponding differences, $1709/gh^3$, calculated from Jeffreys' formula, have also been tabulated for comparison With a depth of 2 mm stability had not been attained when the temperature difference had fallen to 1 3° C To establish a temperature difference of this order uniformly over the whole plate with this depth of chamber did not appear to be possible, and so the critical point for this depth could not be observed accurately

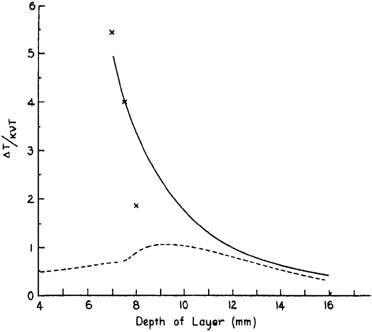


FIG 1—Observed values of $\Delta T/\kappa\nu T$, for different depths, shown by broken curve Continuous curve is for $1709/gh^3$, which represents the corresponding values given by Jeffreys' theory

The observed values of the critical temperature differences necessary to produce motion in the chamber, divided in each case by the mean temperature of the layer and by $\kappa\nu$, are plotted in fig 1, together with corresponding values as computed from Jeffreys' equation The experimental results show fair agreement with the theoretical criterion in layers of depth exceeding 10 mm , but in shallower depths, in which, as will be explained later, the initial motion was different in character from that in deeper layers, the curves in fig 1 deviate widely

3—*Convection patterns*

It was found that the motion in the chamber, as shown by the smoke patterns, varied considerably with the depth of the chamber and with the temperature differences, the results obtained falling into three distinct groups

1—With a depth equal to, or greater than, 10 mm , the smoke patterns are partly in the form of polygonal prismatic cells, in which the motion is inward towards the centre at the top, downward in the centre, and upward along the periphery, and partly in the form of double rolls with descending motion along the boundary separating the two rolls which form a pair Fig 2, Pl 3, shows a photograph taken from above when the chamber has a depth of 10 mm The pattern does not remain fixed, rolls appearing to break up into polygonal cells, and cells to join up into long rolls, in an apparently irregular manner Similar patterns were observed with all depths of chamber from 10 to 16 mm , the diameters of the polygonal cells and the widths of the rolls increasing with increasing depth of the chamber, and also with increasing difference of temperature between top and bottom The boundaries of the smoke patterns showed up more clearly when the difference of temperature between top and bottom of the chamber was increased, but tended to become wavy when this difference became very great Occasionally a circular or elliptical cell was formed out of a polygonal cell, but the polygonal cell is the basic form

A very effective way of demonstrating the nature of the motion in the chamber is to spread a layer of thick smoke gently along the bottom of the chamber After a short interval of time air descending from above punches clean-cut circular holes in the smoke sheet Gradually the air circulating in the cell above the hole tears away tiny wisps of smoke from the surrounding mass of smoke, carrying it upward over the boundary of the hole, then inward towards the centre, and gradually distributing the smoke through the whole of the cell This experiment could only be demonstrated when the difference of temperature in the chamber did not exceed about twice the critical difference shown in Table I

2—When the depth of the chamber is 6 mm or less, the smoke pattern no longer shows any resemblance to polygonal prismatic cells When smoke is now introduced into the chamber it first breaks up into strips parallel to the direction of motion of the smoke into the chamber, as shown in fig 3, Pl 3 In each of these strips there is upward motion along the central line, which appears as a strongly marked white line This motion is opposite to the downward motion along the central line of the double rolls in fig 2, Pl 3 After a brief interval a shallow layer of clear air forms immediately above the heated plate, this being the dust-free space discussed by Aitken (1884), and the clear air bursts upward to the top of the chamber in small vertical columns, which form clear centres at which ascent appears to continue indefinitely The persistence of the ascending motion can be demonstrated by introducing a small quantity of fresh smoke through a capillary tube into the bottom of the chamber immediately below a clear space The fresh smoke immediately rises, and momentarily fills the clear space The centres of convection tend to drift to and fro, so that after a time it is difficult to recognize any definite pattern in the smoke Similar patterns were obtained in shallower depths, but in a chamber of 2 mm depth, with a high difference of temperature, it was found impossible to fill the chamber with smoke

The description of "Type I" is suggested for the motion observed in chambers of depth 10 mm or more, while the motion observed in chambers of depth 6 mm or less may conveniently be described as of "Type II"

3—With depths of chamber less than 10 mm down to 7 mm , it was found that so long as the difference of temperature between top and bottom of the chamber is relatively small, the motion is of Type I, and is similar to that shown in fig 3, Pl 3 With large differences of temperature the motion becomes of Type II, similar to that shown in fig 2, Pl 3, except that the long double rolls shown in fig 2 are now mainly replaced by polygonal cells, with descent at the centre Fig 4, Pl 4, shows this for a depth of 7 mm At the top left-hand corner of the picture, the pattern shown is of Type II, with ascent at the middle of the round structures there shown, probably due to the insufficient heating of the bottom plate in that corner It was noted that when large differences of temperature were set up in a chamber 7 or 7 5 mm in depth, the polygonal cells which then filled the whole chamber showed no tendency either to drift about or to join up into rolls

TABLE II

Depth of chamber mm	Temperature difference ° C
7 0	103
7 5	62
8 0	23

With depths of 6 mm or less, it was found impossible to produce motion of Type I, no matter how great a difference of temperature was set up between the top and bottom of the chamber by heating the bottom plate It was, however, found possible to set up a pattern of Type I in 6 mm by producing a large difference of temperature by cooling the top of the chamber by the evaporation of liquid air

In Table II are given the minimum differences of temperature for the change from convection of Type II to that of Type I These three values are represented in fig 1 by the crosses shown for the corresponding depths It is seen that, particularly for 7 and 7 5 mm , there is excellent agreement between these observations and the theory given by Jeffreys

4—*Comparison with Jeffreys' theory*

The theoretical treatment given by Jeffreys and Rayleigh assumes that the motion is of the type shown in fig 4, Pl 4, though the theory gives no indication of the direction in which the air should circulate in the cellular structures Theoretically either direction of motion might be expected

The results shown in tabular form in Tables I and II may be regarded as experimental confirmation of Jeffreys' theory for all depths down to 7 mm , so far as concerns the temperature difference at which the cellular motion demanded by the theory first appears The experiments show in addition that, in the shallower layers of less than 10 mm , another type of motion, that of Type II, appears at lower temperature differences than are required to produce the cellular convection of the theory

PART II—EFFECT OF SHEAR

1—*Experimental work*

Description of apparatus—For the investigation of the effects of shear the apparatus used in the earlier work was replaced by one in which the top of the chamber was covered by a long glass plate, which could be moved at a variable rate in the direction of its length The apparatus was similar to that used by Graham (1933) for similar experiments, except that the heated metal plate was flanked at its ends by black glass plates, which could be levelled independently so as to lie in the same horizontal plane as the metal plate The arrangements for heating the metal plate were similar to those in the apparatus described in Part I above

The space between the moving glass plate and the fixed metal plate formed the chamber in which the motion was studied The two ends of the chamber were connected by a wooden tunnel, leading into the chamber at one side of each of the black plates The purpose of this tunnel was to provide for the free flow of air when dragged by the upper glass plate, so avoiding return currents in the chamber

Temperature measurements were made by means of two thermocouples mounted close to the upper and lower plates respectively The critical differences of temperature at which motion occurred in this chamber, when the top plate was at rest, were evaluated, and agreed with those given in Table I above Thus the method of measuring temperature in the present series of observations was comparable with that used in the experiments described in Part I

The chamber having been adjusted to the required depth, the metal plate was heated until the difference of temperature between top and bottom of the chamber was of the desired magnitude The top glass plate was arranged so that one end just covered the chamber, and tobacco smoke was introduced into the chamber at the end from which the moving plate travelled, enough smoke being introduced to cover the whole of the black glass plate When the top plate was started, smoke was drawn over the heated metal plate, and the patterns of the motion which it revealed in the part of the chamber above the metal plate were carefully observed

2—*Direct evidence for uniform shear*

A test was made to show that the moving plate produced approximately uniform shear through the whole depth of the chamber This was done by filling the part of the chamber over one of the dark glass plates with smoke, setting the top plate in motion, and observing the behaviour of the smoke through a window cut in the side of the chamber The front of the moving smoke was observed to take the form of an inclined plane, showing that the layer of air in the chamber was subjected to a uniform shear throughout its depth

3—*Effect produced by the boundary between the hot and cold plates*

Since the base of the chamber consisted of three parts, of which the centre one was heated, boundary effects occurred at the edges of the heated plate When the top glass plate was stationary and smoke was blown into the chamber, instability patterns appropriate to the depth of the chamber and the temperature difference between top and bottom formed over the greater part of the metal plate At the boundary between the heated metal plate and the dark glass plates there formed a long roll vortex, the motion of the air in this vortex being upward over the edge of the hot metal plate and downward over the edge of the cooler glass plate If now the top plate was moved, it dragged along with it the boundary vortex, and carried it over the heated plate As soon as the vortex was removed from the boundary, it tended to break up into small circulations appropriate to the depth of the chamber, but still arranged in lines across the plate, while another roll vortex formed on the boundary behind it With the differences of temperature between top and bottom of the chamber set up in the course of these experiments, such transverse rolls were not formed over the central part of the plate, even when the motion of the top plate commenced after the disturbances set up by the introduction of the smoke had died down This suggests that the transverse rolls must be regarded as a boundary effect, due to the nature of the apparatus used, but this cannot be regarded as proved

4—*Effect of shear on motion of Type I*

The effects of shear were first investigated with chamber depths of 7, 8, 12 and 16 mm , with shear ranging from 0 5 to 3 5 cm /sec , and occasionally up to 10 cm /sec , and with varying temperature differences The following results were obtained

a—When the difference of temperature between top and bottom of the chamber was less than the critical value given in Table I, shear produced no smoke pattern, showing that shear did not modify the condition for vertical stability

b—With a difference of temperature that is relatively small but sufficient to produce convection of Type I, and a small rate of shear, transverse rolls were formed at the boundary, and eventually filled the whole chamber The roll vortices were in pairs, adjacent vortices rotating in opposite directions As the rate of shear was increased, the transverse rolls tended to change into longitudinal rolls in the downwind part of the chamber (fig 5, Pl 4)

c—With a relatively large difference of temperature and a small rate of shear, the smoke patterns were generally similar to those obtained with no shear, except that the cellular structures were slightly distorted As the shear was increased, there was noted a marked tendency towards a longitudinal arrangement of these patterns, as shown in fig 6, Pl 5 The greater the temperature difference in the chamber, the greater was the shear necessary to arrange the patterns longitudinally The typical cell showed the form of the one marked *A* in fig 6, Pl 5 Similar patterns were obtained in depths of 8 and 12 mm , but only with high differences of temperature

d—With a high rate of shear, the chamber became filled with longitudinal double rolls, the two rolls in a pair rotating in opposite directions, with descending motion along their common boundary The pattern obtained in 16 mm depth is shown in fig 7, Pl 5 The greater the difference of temperature in the chamber, the higher was the shear necessary in order to produce this effect For example, in a depth of 16 mm , with a difference of temperature of 47° C , no longitudinal rolls were formed by a shear of 2·3 cm /sec , but with a difference of 13° C a shear of 1 3 cm /sec gave well-defined longitudinal rolls A shear of 6 cm /sec was sufficient to produce longitudinal rolls in all the depths investigated, provided the temperature difference did not exceed 100° C With a depth of 16 mm a temperature difference of the order of 100° C combined with a shear of 6 cm /sec gave longitudinal rolls whose edges were wavy, but the waviness disappeared when the rate of shear was still further increased

e—The longitudinal rolls were found to be very stable when once formed, especially in the greater depths investigated They persisted for some time after the top plate was stopped A similar result was derived in the experiments described in Part I above, where it was found that the rolls which formed when smoke was blown into the chamber tended to persist for some time, especially in the greater depths

5—*Effects of shear on motion of Type II*

When patterns of Type II in depths of 6, 4 and 2 mm were subjected to shear, the adjoining columns of clear air joined into lanes roughly parallel to the motion of the top plate, forming ill-defined longitudinal patterns, as shown in fig 8, Pl 6 It will be seen that two lanes of clear air sometimes start at the same point, the smoke between them assuming a lenticular form, with its pointed end facing the direction from which the top plate moves, i e facing upwind

The lengths of the lanes of clear air increase with increasing shear, showing that more and more of the original centres then join up to form a lane of clear air When the rate of shear is small the number of short clear lanes is enormously increased, and the pattern becomes too irregular for any simple description

The boundary phenomenon showed itself by the number of transverse breaks in the patterns near the boundary, as shown in fig 8, Pl 6, on the left-hand side

When the depth of the chamber was reduced to 4 or 2 mm , the patterns observed were precisely similar to those described above, the linear dimensions being reduced in proportion to the depth In none of the depths in which motion of Type II was investigated was any pattern produced by shear when the temperature difference was less than the critical value shown in Table I

6—*Simultaneous production of motions of Types I and II*

Since the phenomena in chambers having depths of 6 and 8 mm differ so widely when the temperature difference is high, an experiment was carried out in which the depth of the chamber varied continuously from 6 mm on one side to 8 mm on the other When the top plate was moved across the chamber, the smoke showed all the usual patterns of Types I and II in the appropriate depths Fig 9, Pl 6, shows the configuration when the shear was stopped

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor D Brunt, at whose suggestion this work was undertaken, for his help during its progress, to Professor G P Thomson for laboratory facilities, to Assistant Professor H S Gregory for help in setting up the apparatus for temperature measurement, and to Messrs Y S Chu and D T E Dassanayake for taking numerous temperature readings

SUMMARY

Part I—A layer of air, having no initial mean motion and enclosed within a chamber, was made unstable by heating from below, and the motion produced, made visible by tobacco smoke, was studied for a range of depths of the chamber and of temperature differences between the top and bottom of the chamber It was found that no motion occurred in the chamber unless the temperature difference exceeded a critical value, which varied with the depth of the chamber

The new results obtained in Part I are

a—That in depths down to 7 mm the temperature difference necessary to produce motion such as was investigated in the Rayleigh-Jeffreys theory is in good accordance with the theoretical expression given by Jeffreys

b—That in all depths below 10 mm a type of motion other than that discussed by Jeffreys occurs with lower differences of temperature than those given by Jeffreys' theory

Part II—When the air in the chamber was subjected to a shear by the motion of the top bounding plate across the chamber, it was found that the

smoke patterns might show transverse rolls, distorted polygons, or longitudinal rolls, the precise forms depending on the magnitudes of the rate of shear and of the temperature differences within the chamber, or that if the original pattern shown in the chamber with no shear were of Type II (characteristic of depths of 6 mm or less), the effect of shear was to give spindle-shaped structures, whose length increased as the rate of shear was increased In shallow depths, 6 mm or less, transverse rolls were never produced, and the structures produced by large rates of shear were not in the form of distinct and well-defined double rolls such as are shown in fig 7, Pl 5 The chief new features in Part II are the measurements of the temperature differences, as well as of shears, associated with the various patterns observed

REFERENCES

Aitken, J 1884 "Collected Works", pp 84–113, *Trans R S Edinb* **32**
Graham, A 1933 *Philos Trans* A **232**, 291–4
Jeffreys, H 1928 *Proc Roy Soc* A, **118**, 195
Rayleigh 1916 "Collected Papers", **6**, 441, *Phil Mag* **32**, 529
Walker, G T 1935 "Clouds, Natural and Artificial" Roy Inst, 8 Feb Lecture published by the Royal Institution

The Production of Neutrons by Bombardment of Beryllium with α-Particles

By T Bjerge

Institute of Theoretical Physics, Copenhagen

(*Communicated by N Bohr For Mem R S—Received* 21 *September* 1937)

1—Introduction

When beryllium is bombarded with α-particles, neutrons are emitted, and this was, in fact, the process giving rise to the discovery of the neutron by Chadwick (1933) The nuclear reaction by which Chadwick explained the production of the neutrons is

$$ {}^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{12}_{6}\mathrm{C} + {}^{1}_{0}n \qquad (1) $$

The energy distribution of the neutrons, when a thick layer of beryllium is bombarded with α-particles from radon, has been investigated by Dunning (1934), who detected the neutrons by recoil protons counted in an ionization chamber connected to a linear amplifier The maximum energy found by Dunning is consistent with (1), inserting the atomic masses now fairly well known from other nuclear reactions and from mass-spectroscopic determinations (Oliphant 1936, Bonner and Brubaker 1936) It appears, however, that in most cases the neutron does not carry away the whole energy available, but leaves the ${}^{12}_{6}\mathrm{C}$ nucleus in an excited state Dunning finds an excited state of c 6^6 e-volts, and Bernardini and Bocciarelli (1936) in similar experiments using polonium α-particles find excited states of $6\ 6^6$, $4\ 5^6$ and $3\ 0^6$ e-volts

This is in good accordance with measurements of Bothe (1936, cf also Maier-Leibnitz 1936), who found that the γ-ray spectrum from beryllium bombarded with α-particles consists mainly of three lines with energies $6\ 7^6$, $4\ 2^6$ and $2\ 7^6$ e-volts The γ-lines found by Crane, Delsasso, Fowler and Lauritsen (1934 and 1935) in other nuclear disintegrations giving rise to the formation of an excited ${}^{12}_{6}\mathrm{C}$ nucleus are consistent with these values, see also the general survey in the report of Livingston and Bethe (1937)

In the experiments just mentioned only neutrons of energies greater than about 1^6 e-volts could be detected When the experimental technique employed for the study of the neutrons emitted from beryllium under

α-particle bombardment permits of measuring also neutrons with energies less than 100,000 e-volts, it appears that there is, besides the neutron groups already discussed, also a large amount of comparatively slow neutrons present This was first pointed out by Auger (1933), who studied the energy distribution by means of a hydrogen-filled Wilson chamber and found that the neutrons could be divided into two main groups, one of them having energies of several million electron-volts (these are the neutrons measured by Dunning and others as already mentioned), the other having energies distributed around a 100,000 e-volts When the neutrons originated from a beryllium-polonium source, the two groups were about equally abundant From a beryllium-radon source, however, the number of neutrons of the slower group was about five times larger than the number of neutrons of the faster group

Further evidence for the existence of a large amount of fairly slow neutrons accompanying the neutron groups of several million electron-volts energies is provided by various experimental facts concerning the slowing down of neutrons to thermal velocities by collisions with hydrogen nuclei Amaldi and others (1935) made a comparison between the density distributions of slow neutrons in a large vessel of water around two different neutron sources, viz a beryllium-radon source and a boron-radon source The distribution around the former showed a much more pronounced accumulation of slow neutrons in the direct neighbourhood of the source than the distribution around the latter, although the faster neutron groups from a beryllium-radon source have rather larger energies than the neutrons from a boron-radon source Also the total number of neutrons emitted from a beryllium-radon source can be determined in various ways by slowing them down in hydrogenous media Paneth, Gluckauf and Loleit (1936) find this number to be greater than 7000/sec /millicurie of radon, and Amaldi and Fermi (1936) find *c* 28,000 (the figure now generally adopted is about 20,000), whereas for the neutrons of more than about 1^6 e-volts Dunning (1934) finds a yield of *c* 1500/sec /millicurie of radon Jackel (1934) finds 9000 for the number with energies higher than 90,000 e-volts A comparison between these figures shows that only about 10 % of all the neutrons from a beryllium-radon source belong to the faster group, while the rest have energies of the order of 100,000 e-volts or less

2—THEORETICAL CONSIDERATIONS

Starting from the point of view put forward by Bohr (1936) the nuclear reaction in question has to be considered in two steps First the α-particle and the ${}^{9}_{4}\mathrm{Be}$ nucleus fuse together, thus forming a compound nucleus ${}^{13}_{6}\mathrm{C}$ of high energy content

$$ {}^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{13}_{6}\mathrm{C} \tag{2} $$

Then this highly excited nucleus may disintegrate in a number of different ways, each of which has its own probability

Thus the whole of the excitation energy ($12\text{–}14^{6}$ e-volts) may at some time become concentrated on one neutron, giving rise to the disintegration

$$ {}^{13}_{6}\mathrm{C} \rightarrow {}^{12}_{6}\mathrm{C} + {}^{1}_{0}n, \tag{3} $$

where the ${}^{12}_{6}\mathrm{C}$ nucleus is left in the ground state, but it is more probable that only so much energy will be concentrated on one neutron, that the ${}^{12}_{6}\mathrm{C}$ nucleus is left in one of its excited states From more detailed considerations (Bohr 1937, cf Bohr and Kalckar 1937 and Weisskopf 1937) it follows, in fact, that the most probable values for the kinetic energy of the ejected neutron should be about 1 or 2^{6} e-volts for the disintegration in question, and that the number of neutrons ejected with greater energies should decrease exponentially with increasing energy The distribution of fast neutrons from beryllium bombarded with α-particles, as found experimentally, e g by Bernardini and Bocciarelli (1936), is in agreement with these expectations

The ejection of a neutron with an energy as small as 100,000 e-volts, on the other hand, should be a very unlikely process, and the presence of a large number of such fairly slow neutrons can, therefore, not be explained by the reaction (3) It is, however, quite likely, and indeed highly probable, that in addition to (3) the disintegration

$$ {}^{13}_{6}\mathrm{C} \rightarrow {}^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \tag{4} $$

takes place If in this process the ${}^{9}_{4}\mathrm{Be}$ nucleus is left in the ground state, the total process ((2) and (4)) is a scattering of the α-particle If, however, the ${}^{9}_{4}\mathrm{Be}$ is left in an excited state (which on Bohr's theory would be more probable, as soon as the energy is sufficient), it would, on account of its small binding energy, have a great chance to explode

$$ {}^{9}_{4}\mathrm{Be} \rightarrow {}^{8}_{4}\mathrm{Be} + {}^{1}_{0}n, \tag{5} $$

or more likely

$$ {}^{9}_{4}\mathrm{Be} \rightarrow 2{}^{4}_{2}\mathrm{He} + {}^{1}_{0}n \tag{6} $$

In (6) especially, the energy of the neutron might be quite small, because the mutual repulsion of the two α-particles will consume a large part of the energy present.

The total of (4) and (6) is a process by which an excited $^{13}_{6}C$ nucleus breaks up into three α-particles and a neutron, such a process has been observed by Cockcroft and Lewis (1936) and by Bonner and Brubaker (1936) bombarding boron with deuterons

$$^{11}_{5}B + ^{2}_{1}H \to ^{13}_{6}C \to 3^{4}_{2}He + ^{1}_{0}n \tag{7}$$

Chadwick, Feather and Davies (1934), bombarding carbon by neutrons in the expansion chamber, have found a triple fork of tracks presumably caused by the disintegration

$$^{12}_{6}C + ^{1}_{0}n \to ^{13}_{6}C \to 3^{4}_{2}He + ^{1}_{0}n \tag{8}$$

An excited $^{8}_{4}Be$ nucleus, breaking up in accordance with (6), can also be formed by the process

$$^{7}_{3}Li + ^{2}_{1}H \to ^{9}_{4}Be \to 2^{4}_{2}He + ^{1}_{0}n, \tag{9}$$

see e g Oliphant, Kempton and Rutherford (1935)

The interpretation of the slow-neutron group expressed by (2), (4) and (6) has been proposed already in a very similar form by Auger (1933), who noticed that the different behaviour of the two groups, when the energy of the incident α-particles is changed, makes it likely that they are connected with two different modes of disintegration With the better knowledge of atomic masses which we have now, it is possible to give this argument the following more precise form whereas (1) is energetically possible for any energy of the incident α-particle and therefore may take place as soon as the α-particle has sufficient energy to break through the potential barrier, the process consisting of (2), (4) and (6) requires that the α-particle must deliver more than 2^{6} e-volts to the system in addition to the energy necessary for an α-particle to break out through the potential barrier The hypothesis, that the fast neutron group is due to (1), the slow neutron group to (2), (4) and (6), is therefore supported by the fact found by Auger, that the slow group becomes relatively more abundant, when more energetic α-particles are used

3—EXCITATION CURVES FOR NEUTRONS

The experimental work reported in the present paper was undertaken at the suggestion of Professor Bohr with the object of obtaining more evidence as to the origin of the slow-neutron group mentioned above

If the α-particles emitted in the processes (4) and (6) could be detected and studied either in an expansion chamber or by counting methods, this would no doubt give the most direct information Unfortunately, however, the energies of these α-particles are small and it would be difficult to detect them, especially as the α-particles scattered without any great energy loss would probably predominate As this method, therefore, did not seem very promising I confined myself to trying to establish more quantitatively the different behaviour of the two main groups of neutrons when the energy of the incident α-particles is varied

Excitation curves, giving the yield of fast neutrons from beryllium as a function of the energy of the incident α-particles, have been obtained by

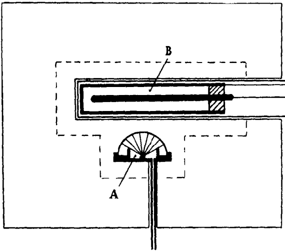


Fig 1—Arrangement for measuring the yield of neutrons as a function of the α particle energy

Chadwick (1933), who counted the neutrons by the recoils in an ionization chamber connected to a linear amplifier, and by Rasetti (1932) and Bernardini (1933), who measured the ionization current in a methane pressure ionization chamber The curves of Chadwick and Bernardini are reproduced in fig 2 A sensitive method for measuring the yield of slow as well as fast neutrons is to surround the neutron source with paraffin and detect the neutrons, slowed down by collisions with the hydrogen nuclei of the paraffin, by means of a boron-lined ionization chamber connected to a linear amplifier, or still better by an ionization chamber filled with BF_3 at a suitable pressure The experimental arrangement is shown in fig 1 The neutron source *A* consists of a copper hemisphere of 2 5 cm radius, lined inside with a beryllium layer of 4 mm air equivalent (This layer was evaporated on to the copper hemisphere by placing a beryllium crystal (c 50 mg) in the centre and heating it by bombardment *in vacuo* with electrons (from a hot tungsten filament) accelerated through some thousand volts) At the centre of the hemisphere is placed a polonium source (about 12–14 millicuries) deposited on a 7 mm silver disk The hemisphere together with the brass plate holding the polonium source form an airtight box

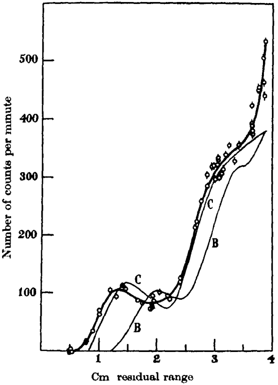


FIG 2—Yield of neutrons as a function of α particle energy, large paraffin block (*C* and *B* are the curves of Chadwick and of Bernardini, respectively)

A given energy of the α-particles falling on the beryllium is obtained by filling this box with carbon dioxide at a given pressure The α-particles leaving the surface of the source at nearly glancing angles are cut off by a cylindrical metal screen which allows only the rays within a cone of c 120° to reach the beryllium

The ionization chamber *B* is a steel cylinder, 2 5 cm inner diameter and c 15 cm long, filled with boron trifluoride at a pressure of about 1 5 atm The inner electrode is connected to a Wynn-Williams linear amplifier with a mechanical counter driven by a thyratron The steel cylinder has a positive potential of *c* 1000 V, it is surrounded by an earthed brass cylinder of 4 cm outer diameter The chamber was made by Dr Jørgen Koch, who kindly put it at my disposal

In one experiment *A* and *B* were embedded in a block of paraffin wax 35 × 30 × 24 cm , in another, they were surrounded by a layer of paraffin wax, 2 cm thick on an average

The result of the former experiment is plotted in fig 2 To obtain each point counts were taken for a period of 10–30 min The values of the ordinates have been corrected for the time-resolving power of the mechanical counter, and the "natural effect" of the chamber (15/min) has been subtracted The residual range is taken as 3 90 cm less the air equivalent of 2 5 cm CO_2 of the given pressure (and temperature) The errors indicated are the statistical mean-square errors, the curve was taken in both directions several times, and every time the steep rise beginning at about 3 5 cm residual range was conspicuous The sensitivity of the detecting apparatus, however, appeared to vary a little from day to day, so that, when plotting all the points together, their spread around the best curve through the points is somewhat larger than to be expected from the statistical errors For comparison, the curves obtained by Chadwick and by Bernardini are also given in the figure All three curves have a very similar appearance except that

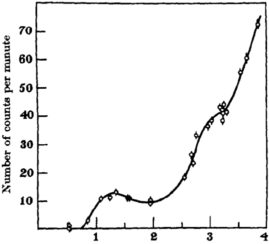


Fig 3—Yield of neutrons as a function of α particle energy, small layer of paraffin

the fast-neutron curve of Chadwick does not show the steep rise beginning at c 3 5 cm residual range This rise therefore is probably due to a group of fairly slow neutrons

In order to establish this point more clearly, the large paraffin block was now taken away, and the source and counting chamber were surrounded with a 2 cm layer of paraffin wax Fast neutrons of several million electron-volts energy have a greater chance of escaping right through this layer than have neutrons of c 100,000 e-volts The measured yield should therefore be more reduced for the faster group than for the slower, when changing to the smaller paraffin layer The result is plotted in fig 3, it will be seen that the group starting at 3 3 cm * is relatively stronger than in fig 2, whereas the curve for smaller energies of the α-particles has a shape practically identical with that of fig 2 This is seen very clearly in fig 7, where the two curves are scaled up to have the same ordinate at 3 2 cm It may be mentioned that a curve taken with an intermediate amount of paraffin had an intermediate shape

4—EXCITATION CURVES FOR γ-RAYS

If the origin of the slow-neutron group is that mentioned in § 2, clearly no γ-rays will be emitted in connexion with the emission of the neutrons of this group The excitation curve for γ-rays from beryllium bombarded with α-particles has been measured by Becker and Bothe (1932) Their curve, which is reproduced in fig 6 (B & B), is derived from an integral curve obtained with a thick beryllium layer, and it is difficult to obtain the details correctly by this method I therefore considered it of interest to measure this excitation curve, using the same neutron source as was used in the experiments of § 3

Fig 4 shows the experimental arrangement *A* is the same neutron source as in fig 1 The γ-rays emitted from the beryllium layer produce Compton electrons in the copper hemisphere and the 1 mm copper plate *P* Some of these electrons will penetrate the two thin walled Geiger counters *B* and *C*, of which *B* is shown in section They have 3 cm diameter and 4 cm inner length and are made of 0 1 mm copper foil with ebonite bungs and a 0 05 mm tungsten wire at the axis The cylindrical wall is strengthened by ring-shaped corrugations The gas-filling is dry hydrogen at 100 mm pressure The two Geiger counters are connected to two identical two-stage amplifiers, from which the impulses are fed on to two triodes in a Rossi coincidence circuit When both Rossi valves are extinguished simultaneously, the impulse in their common anode circuit strikes a thyratron, which drives a mechanical counter The high tension for the Geiger-counter cylinders (1500 V negative) was taken from one dry-cell battery, interaction between them being prevented by filter circuits and by the earthed screen *S* The source and counters were arranged in a horizontal plane in order to minimize the number of coincidences from cosmic particles No shielding was provided, because neutrons captured in the shielding material might have produced γ-rays sufficiently hard to give coincidences The screen *S* was a 0 4 mm iron plate, which means that the total stopping power of the material between the two counters was equivalent to 2 mm aluminium, this will cut off completely any secondary electrons produced by the γ-rays

---

* It is difficult to say whether the apparent difference in abscissa from 3 5 to 3 3 cm is real or is due to experimental errors

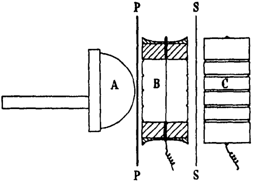


FIG 4—Measurement of γ rays by coincidences

of the polonium (Bothe and Becker 1930) The "natural effect" of each counter was c 120/min , the time in which the Rossi valve is extinguished by any impulse in its Geiger counter is c 2×10^{-4} sec (calculated from the time constants of the circuits), from which it is found that the accidental coincidences should be about five per hour, in good agreement with the experimental value six or seven

When the source is pumped out so that the α-particles fall on the beryllium with full energy, the number of counts per hour increases to an amount between 60 and 100 This number varied slowly during the two weeks in which the measurements were made, the variation seemed to follow the variation of the moisture in the room

When plates of aluminium were interposed between the counters, the number of coincidences decreased as shown in the curve, fig 5 This curve is a measure of the hardness of the γ-rays and is practically identical with the one given by Becker and Bothe (1932)

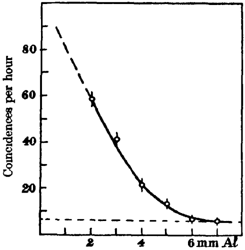


FIG 5—Number of coincidences plotted against thickness of aluminium between counters

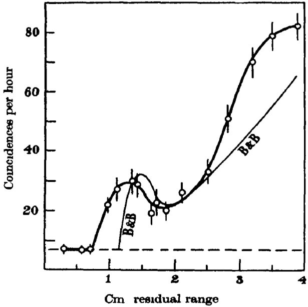


FIG 6—Yield of γ rays as a function of the α particle energy. "B & B" is the curve of Becker and Bothe

When only the iron plate S was interposed between the counters, and the CO_2 pressure in the neutron source was varied, the curve of fig 6 was obtained The value at 3 9 cm residual range (all CO_2 pumped out) was used as a standard of reference in the following way For every measurement the reference point was first counted through 1 hr , then the point in question was counted through 1 hr , then again the reference point Usually the measurement was repeated In this way two or three points were taken during a day and one during the night (A similar procedure was used for the points of fig 5 The values for the reference point varied very slowly and steadily between 80 and 100/hr when the curve fig 6 was taken, and was about 60 during the measurement of fig 5)

5—CONCLUSION

In fig 7 the three excitation curves of figs 2, 3 and 6 are plotted together on an arbitrary scale of ordinates, chosen so that they coincide at the point with abscissa 3 2 cm residual range The yield of faster neutrons follows the yield of γ-rays, whereas the group of slower neutrons emitted when the incident α-particles have a residual range of more than 3 4 cm (average of the two values), is not accompanied by γ-rays

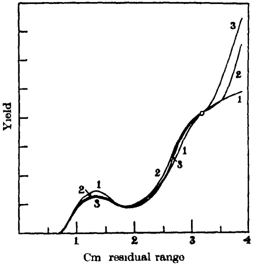


FIG 7—Excitation curves for beryllium bombarded with Po α particles 1, γ rays, 2, neutrons, large paraffin block, 3, neutrons, small paraffin layer

While the faster groups and the γ-rays are accounted for by the process (1), this process cannot, on theoretical reasons, account for the slower group The only natural explanation of this group is that given by the reactions (2), (4) and (6) The present experiments support this explanation by showing, in the first place, that the slower group of neutrons is not accompanied by γ-rays and, in the second place, that this group is emitted as soon as the α-particles have a range of at least 3 4 cm , corresponding to an energy of 4 9^6 e-volts Within the limits of error this is just the energy, which the α-particle must have in (2), in order that the $^9_4$Be left in (4) has sufficient energy to split up into two α-particles and a neutron (process (6)) This energy is calculated as follows of the energy of the incident α-particle (process (2)) only $\frac{9}{13}$ is changed into inner energy of the $^{13}_6$C nucleus When this nucleus splits up (process (4)), the $^4_2$He takes away a kinetic energy of at least 0 7^6 e-volts (measured in the centre-of-gravity system), and the $^9_4$Be at least 0 3^6 e-volts [This is deduced from the fact derived from fig 7, that an α-particle must have a kinetic energy of 1 4^6 e-volts (0 8 cm range) in order to have a reasonable chance of penetrating into a beryllium nucleus When measured in the centre-of-gravity system, this energy becomes 0 7^6 e-volts, and the kinetic energy of the $^9_4$Be nucleus before the collision becomes 0 3^6 e-volts If we suppose the corresponding mutual velocity to be necessary for penetrating the potential barrier, this will also be the minimum mutual velocity, when a $^{13}_6$C nucleus splits up into $^4_2\text{He}+^9_4\text{Be}$] Further, when in process (6) the $^9_4$Be splits up into a neutron and two α-particles, these latter will have a total kinetic energy (in their centre-of-gravity system) of at least *c* 0 5^6 e-volts (the potential energy of an α-particle in the Coulomb field of $^4_2$He being half the energy in the field of $^9_4$Be at the same nuclear distance) Taking the masses given by Oliphant (1936), we thus find for the excited $^9_4$Be (process (6)) the mass

$$2\times 4\ 0039+1\ 0091+0\ 0005 = 9\ 0174,$$

i e 2 3^6 e-volts more energy than the normal $^9_4\text{Be} = 9\ 0149$ In order to give the $^9_4$Be this excitation and in addition provide the 1 0^6 e-volts necessary for the splitting up of $^{13}_6$C in process (4), the incident α-particle must have an energy of $\frac{13}{9}\ (2\ 3+1\ 0) = 4\ 8^6$ e-volts

The experimental value 4 9^6 e-volts has some uncertainty (*c* 0 2^6 e-volts), especially because the finite thickness of the beryllium layer may have changed the curve slightly in a way that is difficult to estimate A similar uncertainty has to be ascribed to the calculated figure 4 8^6 e-volts

I wish to thank Professor N Bohr for his kind interest in the work and Mr F Kalckar for instructive discussions My thanks are also due to Mr E Kofoed and Mr K Overgaard for their valuable assistance with the experiments and to Dr J C Jacobsen for helpful advice on the coincidence counting circuit

6—SUMMARY

In the energy spectrum of the neutrons from beryllium bombarded with α-particles there are present some groups with energies of several million electron-volts These are accounted for by the process

$$ {}^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{12}_{6}\mathrm{C} + {}^{1}_{0}n \qquad (1) $$

The ${}^{12}_{6}\mathrm{C}$ may be left in an excited state and afterwards emit a γ-quantum

There is also, as shown by Auger and others, a neutron group of energies of the order of *c* 100,000 e-volts Auger gave reasons for supposing that these neutrons are created in a process, where the incident α-particle first transfers some energy to the ${}^{9}_{4}\mathrm{Be}$, which then disintegrates in the following way

$$ {}^{9}_{4}\mathrm{Be} \rightarrow 2{}^{4}_{2}\mathrm{He} + {}^{1}_{0}n \qquad (6) $$

This view is supported by the ideas of Bohr on nuclear reactions The present paper describes experiments which give the yield of the faster and slower neutron group and of the γ-rays as a function of the α-particle energy It is shown that the slower group is not accompanied by γ-rays, and that it begins to be emitted at the α-particle energy which is just necessary for the process (6) to take place

[*Note added in proof*—In collaboration with Mr K Overgaard an attempt has been made to extend the excitation curve for neutrons to higher α-particle energies by using thorium (B + C) preparations of 0 3–0 5 millicuries When the range of the incident α-particles increases from 3 9 to 8 6 cm the yield of neutrons increases by a factor of 3, the rate of increase being apparently fairly steady with a tendency to diminish as was to be expected Still, this can only be taken as a rather rough measurement, the effect obtained being too small for ascertaining the details of the curve]

References

Amaldi, d'Agostino, Fermi, Pontecorvo and Segré 1935 *Ricerca Scientifica*, **6**, 1, nos 11–12

Amaldi and Fermi 1936 *Phys Rev* **50**, 899

Auger 1933 *J Phys Radium*, **4**, 719

Becker and Bothe 1932 *Z Phys* **76**, 421

Bernardini 1933 *Z Phys* **85**, 555

Bernardini and Bocciarelli 1936 *Accad Lincei Atti*, **24**, 59 and 132

Bohr 1936 *Nature, Lond*, **137**, 344

— 1937 *Science*, **86**, 161

Bohr and Kalckar 1937 *K danske vidensk Selsk Skr* **14**, no 10

Bonner and Brubaker 1936 *Phys Rev* **50**, 308

Bothe 1936 *Z Phys* **100**, 273

Bothe and Becker 1930 *Z Phys* **66**, 307

Chadwick 1933 *Proc Roy Soc* A, **142**, 1

Chadwick, Feather and Davies 1934 *Proc Camb Phil Soc* **30**, 357

Cockcroft and Lewis 1936 *Proc Roy Soc* A, **154**, 246 and 261

Crane, Delsasso, Fowler and Lauritsen 1934 *Phys Rev* **46**, 1109

— — — — 1935 *Phys Rev* **48**, 100

Dunning 1934 *Phys Rev* **45**, 586

Jäckel 1934 *Z Phys* **91**, 493

Livingston and Bethe 1937 *Rev Mod Phys* **9**, 245

Maier-Leibnitz 1936 *Z Phys* **101**, 478

Oliphant 1936 *Nature, Lond*, **137**, 396

Oliphant, Kempton and Rutherford 1935 *Proc Roy Soc* A, **149**, 406

Paneth, Gluckauf and Loleit 1936 *Proc Roy Soc* A, **157**, 412

Rasetti 1932 *Z Phys* **78**, 165

Weisskopf 1937 *Phys Rev* **52**, 295

On the Penetrating Component of Cosmic Radiation

By H J Bhabha

Gonville and Caius College, Cambridge

(*Communicated by R H Fowler, F R S —Received* 4 *October* 1937)

Introduction

The position in cosmic radiation has changed considerably in the last year both from the experimental and the theoretical side, so that it is now possible to co-ordinate the various independent observations to a degree which was not hitherto possible and to draw some important conclusions from them Since it has been shown by Rossi (1934*a*) and his co-workers, Auger and Ehrenfest (1934) and Street, Woodward and Stevenson (1935) that there are single ionizing particles in cosmic radiation which penetrate more than a metre of lead, and further, according to the theory, no electron of any reasonable energy can make its effects felt through more than about 15 cm of lead, it has become clear that the behaviour of the penetrating component of cosmic radiation faces us with at least one of the two following conclusions

a—The theoretical formulae for the energy loss of fast electrons break down for energies above some critical energy, where this critical energy may or may not depend on the material

b—The penetrating component does not consist of electrons

Moreover, since Blackett (1937*a*) has found that there is a rough equality in the number of positive and negative particles up to the highest measurable energies, it follows that the second alternative already demands the existence of a hitherto unknown particle, since even if we assume, as was supposed at first, that these penetrating particles are protons, for which the radiation loss is small due to their larger mass, resulting in a corresponding increase of penetrating power, the presence in the penetrating group of negatively charged particles forces one to admit the existence of negative protons. The assumption that these particles are protons has, however, met with the difficulty noticed by several investigators that far fewer protons, identifiable at the end of their range by a heavy track in a Wilson Chamber, are observed than there should be, a difficulty which can only be removed by certain plausible but *ad hoc* assumptions about processes (such as nuclear collisions and explosions) which would remove protons from the beam before they became slow enough to show a noticeably greater ionization

Street and Stevenson (1937) have, however, shown that among those particles which have already traversed 10 cm of lead, there are some with curvatures which would correspond to energies of less than $7\ 5 \times 10^8$ e-volts if they were electrons, whose behaviour is not that to be expected theoretically for electrons, and which are certainly not protons The latest experiments of Blackett and Wilson (1937) also show that at energies as low as 4×10^8 e-volts in lead, cosmic-ray particles have an energy loss much less than the theoretical loss for electrons, and they further claim that it is possible to exclude the hypothesis of protons since they would show a noticeably greater ionization along their tracks at these energies For the further discussion, therefore, we will only consider the conclusion (*b*) stated above in the form

b—There are in the penetrating component of cosmic radiation new particles of electronic charge of both signs, and mass or masses intermediate between those of the electron and proton For brevity, in the further discussions we shall describe such particles simply as heavy electrons

The two alternatives (*a*) and (*b*) are not mutually exclusive

As has been shown by Rossi (1934*b*) and especially stressed by Auger and Leprince-Ringuet (1934), the radiation at and above sea-level definitely consists of two groups distinguishable by their penetrating power If we measure the absorption in lead of the vertically incident particles at sea-level, we find that the first 10 cm of lead absorb 25–30 % of the total number of particles At about this point a distinct change in the slope of the absorption curve occurs, and a metre of lead only serves to decrease further the number of particles by about 30 % We shall call the group of particles absorbed in about 10 cm of lead the soft group, and the other the hard or penetrating group It should be emphasized that in this paper the words soft and hard are used to denote only the penetrating power of the particles and have nothing whatsoever to do with their energy

As we shall see in the next section the soft component is not in equilibrium with the hard component above sea-level Further, the hard component obeys the mass absorption law, whereas the soft component shows an absorption per atom which varies as the square of the atomic number as would be the case for electrons obeying the theory

In recent papers it has been shown independently by Carlson and Oppenheimer (1937) and Bhabha and Heitler (1937) that a consistent application of the formulae of Bethe and Heitler (1934) to the successive emission of radiation by electrons and the creation of electron pairs by

γ-rays cannot only explain the showers and bursts observed in cosmic-ray phenomena, but is capable of describing the typical features of these phenomena, namely, the transition curves investigated by Rossi (1933 *b*) and the absorption curve in the atmosphere It is therefore certain that the soft group consists of electrons and positrons and their accompanying γ-radiation, and its behaviour is described correctly by the theory

In view of this fact it is of interest to discuss the experimental material, particularly with regard to the hard component, and to see whether it supports the first or the second of the alternatives mentioned above We carry through this discussion in § 1, where it will appear that although assumption (*a*) may also be true *for extremely high energies*, it is *insufficient* to explain all the facts, and that there are sufficient grounds to justify us in considering the presence of new particles of electronic charge and mass between those of the electron and proton at least as a possibility Under these circumstances we think it not unprofitable to consider the behaviour of such particles theoretically, in order that a comparison may then be made with experiment Of course only that part of the behaviour of such a particle can be calculated which depends essentially only on its charge and mass, namely, the magnitudes of the various collision processes and the ionization and radiation losses, as also the creation of pairs of such particles by γ-rays This has been done in § 2, assuming that the particle obeys the Dirac equation We shall see that the radiation loss is not just inversely proportional to the square of the mass of the particle, since for particles of different masses the effect of screening is different Moreover, since we do not know what mass to attribute to such a particle if it exists, we have carried through the calculations for particles of rest energy equal to 5×10^6 e-volts and 5×10^7 e-volts, i e 10 and 100 times the electron mass respectively, so that together with the already known results for the electron and proton, we obtain curves from which the energy loss can be read off for particles of any given rest mass and energy It will appear in the course of this paper that if the existence of new particles is to be assumed as a solution of the various difficulties which still remain in the cosmic-ray phenomena, it seems probable that particles of various different rest masses will have to be assumed with perhaps the possibility of particles making transitions from one rest mass to another

In § 4 we discuss the production of showers by heavy particles, that is, particles with any mass greater than that of the electron, either by the emission of a sufficiently large quantum of radiation, or by the production of a sufficiently fast secondary electron by collision In § 3 we calculate the average number of positive and negative electrons accompanying the penetrating heavy particles as a result of these processes, thus forming a soft component in equilibrium with the penetrating particles In § 5 we have briefly considered the effect of the creation of pairs of heavy particles in decreasing the rate at which cascade processes die out This may find a possible application in the penetration of particles to sea-level

1—DISCUSSION OF THE EXPERIMENTAL MATERIAL

A—Latitude effect

We have already stated in the introduction that the soft group (defined as those particles absorbed in about the first 10 cm of lead) is not in equilibrium with the hard group This statement is established by the experiments of Auger, Ehrenfest and Leprince-Ringuet (1936), who have measured the absorption curve in lead at sea-level and at Jungfraujoch Table I gives their results

TABLE I

	Jungfraujoch	Sea level	Absorption coefficient in air in cm^2/g	Absorption coefficient in lead in cm^2/g
Hard group	190	120	$0\ 70 \times 10^{-3}$	$0\ 70 \times 10^{-3}$
Soft group	170	25–30	6×10^{-3}	$32 \pm 2 \times 10^{-3}$

The first two columns give in arbitrary units the number of particles of each group found at the two heights, the third column then gives the absorption coefficient deduced from these figures assuming an exponential absorption The last column gives the absorption coefficients deduced from the absorption curve in lead These results again confirm those of Rossi, Alocco (1935), and of Clay (1936), and show that the hard component obeys a mass absorption law The soft group, on the other hand, is seen to show an absorption per atom which is roughly proportional to the square of the atomic number

The above figures show that the soft group increases with height much more rapidly than the hard group Ionization measurement by Compton and Stevenson (1934) and Bowen, Millikan and Neher (1934) at great heights with an ionization chamber shielded by 6 and 12 cm of lead confirm this, and show that at these heights the hard component contributes less than 30 % to the total ionization The shower intensity as measured by Woodward (1936) and Braddick and Gilbert (1936) increases with height much more rapidly than the total radiation and runs more parallel with the intensity of the soft group Heitler (1937) has shown that the variation of showers with altitude and latitude can be understood if they are due predominantly to the soft component Further, Montgomery and Montgomery (1935) find that the intensity of bursts increases with height even more rapidly than the shower intensity, and recently Young (1937) has shown for small bursts containing ten particles and more that the increase with height of bursts of a given size is greater, the larger the size of the bursts Thus we regard this as evidence that bursts, which have been shown by Ehrenberg (1936) to be very large showers, are also predominantly due to the soft component But such a view is incompatible with the assumption (*a*) For on this assumption, the soft component consisting of particles which are absorbed in about 10 cm of lead must be electrons of energy below the critical energy in *lead*, which is less than 4×10^8 e-volts Such electrons would be incapable of producing bursts, merely because they have less energy than the bursts themselves There is no difficulty of this sort on assumption (*b*), since if the soft component consists of electrons obeying the theory, they would be absorbed in about 10 or 15 cm of lead even for energies of the order of 10^{11} e-volts

We finally come to consider the latitude effect at sea-level Bhabha and Heitler (1937) have shown that the shape of the absorption curve in the atmosphere is a proof that electrons of energy near 3×10^9 e-volts (the lowest energy which can reach a magnetic latitude 50° due to the magnetic field of the earth) multiply according to the cascade theory Heitler (1937) has further demonstrated by a more detailed analysis of the same curve that if a breakdown of the theoretical energy loss formulae takes place, the breakdown energy must be at least as high as 5×10^9 e-volts The latest measurements of Bowen, Millikan and Neher (1937) show that particles of energy at least as high as 10^{10} e-volts produce a rapid multiplication in the upper layers of the atmosphere which is at least in qualitative agreement with the cascade theory, thus proving that the radiation loss of electrons demanded by quantum mechanics is correct in air for energies up to 10^{10} e-volts

Let us consider these facts on assumption (*a*) Now the theory shows (Bhabha and Heitler 1937, p 454) that the chance of an electron of energy less than 10^{10} e-volts making its effect felt at sea-level is negligible Further, particles of energy greater than 10^{10} e-volts will not show a latitude effect at latitudes greater than about 35° Thus there should be no variation of intensity at sea-level at latitudes greater than 35°

Bhabha and Heitler have already shown that if the theory of energy loss for electrons be right for all energies, then no latitude effect at sea-level could be due to the electrons Our present considerations show that even if the theory of energy loss break down above some critical energy, no latitude effect beyond 35° could be due to electrons, provided the breakdown energy is greater than 10^{10} e-volts Since at sea-level the latitude effect starts at about 50°, and is quite considerable at 35°, we regard this as a very strong argument against the hypothesis (*a*) and in favour of the existence of new particles

Indeed, in our opinion, the very discrepancy between the theoretical absorption curve in the atmosphere and the experimental difference curve of Bowen, Millikan and Neher for primaries of an average energy of 10^{10} e-volts is evidence of the existence of new particles For the first part of the experimental curve shows that an enormous multiplication takes place in the atmosphere, thus establishing the existence of large radiative losses, from which it follows merely from arguments of self-consistency that the rest of the curve should agree with the theoretical curve at least as regards the order of magnitude, whereas in fact some thirty times as many particles are found at sea-level as there should be We believe that a large fraction of these sea-level particles are heavy electrons, either of primary origin, or secondaries created in the atmosphere, together with the electron component in equilibrium with them which they produce

B—Bursts and transitions curves

It has been shown by Bhabha and Heitler, and Carlson and Oppenheimer in the papers quoted above that not only does the theory show that electrons and γ-rays will create showers in their passage through various substances, but that it also gives the shape of the curve connecting the number of showers containing a given number of particles as a function of the thickness of the material in which they are produced These curves are like the curves found by Rossi (1933) for the number of coincidences between a number of counters placed below a plate of some heavy substance, usually lead, plotted as a function of the thickness of the plate Further, Bhabha and Heitler have shown that the maximum of the Rossi curve for large showers should lie at greater thicknesses than the maximum of the Rossi curve for smaller showers, a prediction which has been confirmed recently by Auger, Ehrenfest, Freon and Grivet (1937) for small showers containing a few particles Moreover, Bøggild (1936) has also observed that for bursts of different size there are indications of the same shift of the maximum

The questions we have to answer then are (1) Do electrons of sufficiently high energy produce bursts by cascade multiplication only, as the quantum theory predicts? (2) Do electrons and other particles produce showers in *one elementary process*? The two alternatives are not mutually exclusive

Before we discuss the bursts, we will analyse the Rossi curve for showers in some detail The Rossi curve for showers at and above sea-level as measured by Auger and Meyer (1933) has the shape shown by curve 1, fig 1 It is important to notice that after the maximum the curve falls away rapidly till about 10 cm of lead, after which the decrease would be much more gradual, being then comparable to the decrease of the penetrating component (cf 4) The same curve, measured by these authors at a depth below the ground equivalent to 30 m of water, where only the hard component is found, has the shape shown by curve 2, and below 75 m by curve 3 The decrease of both these curves after the maximum has an absorption coefficient ten times less

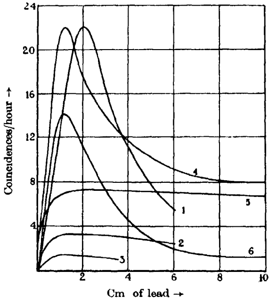


FIG 1—Rossi curves for showers Curves 1, 2, 3, Auger and Meyer, curves 4, 5, 6, Schwegler The ordinates of curves 1, 2, and 3 have been reduced so that the maximum heights of 1 and 4 shall be the same

than that of (1) and is parallel to the decrease of the vertical intensity Moreover, the observation of Clay, Gemert, and Wiersma (1936) that after thicknesses of 200 g /cm ^{2} the decrease of the number of showers is parallel to the decrease of the primary intensity also shows that at least some showers are to be associated with the penetrating component It is therefore reasonable to suppose that the sea-level curve 1 is made up of the superposition of two effects, those due to the soft and hard components respectively This supposition has been confirmed directly by Schwegler (1935) By placing a 10 cm lead block between the three counters used for recording the triple coincidences and thus eliminating the effect of the soft component he gets curve 5, being the number of coincidences plotted as function of the lead above the counters Without the 10 cm block he gets the usual curve 4 The difference between the two is shown by the curve 6, and gives then the showers produced by the soft component alone This is just what we should expect theoretically, for 10 or 15 cm of lead should suffice to absorb more or less completely all electrons or γ-rays of energy below 10^{11} e-volts

The Rossi curve for bursts measured by Bøggild (1936) in iron and Carmichael (1936) in lead have the same shape as curve 1, except that the maximum lies at greater thicknesses We wish to emphasize particularly the fact that after the maximum the *decrease is much more rapid than that of the total intensity* Moreover, Bøggild has found that there are definite indications that the maximum shifts to greater thicknesses for larger bursts We now wish to consider these facts in the light of the assumptions (*a*) and (*b*)

We will first discuss them on the assumption that for energies above a certain critical energy the theory of energy loss fails, and that above these energies the radiation loss is much less than the theoretical, and goes more or less gradually to zero with increasing energy According to Blackett and Wilson's experiments this critical energy must be below 4×10^8 e-volts in lead The energy of the particle starting the burst must be at least as great as the total energy of the burst, which in the case of a burst of 100 particles of 5×10^6 e-volts each amounts to 5×10^8 e-volts and is higher than the critical breakdown energy Since, according to the cascade theory, bursts of a hundred particles or more require an energy in lead which is at least $1\ 5 \times 10^9$ e-volts, it means that the cascade theory of bursts will fail A burst would then have to be started by some elementary process in which a number of very energetic particles or quanta are created at once Of course, those electrons or quanta so formed, whose energy was less than the critical energy, or became less than it on penetrating the lead plate, would then have the theoretically correct energy loss, and multiply according to the cascade theory

But the Rossi curve for such bursts will show a shape characteristically different from curve 1 As the thickness of the lead plate above the ionization chamber is increased from zero, bursts started anywhere in the plate will be registered, and so the number of bursts will increase The effect of adding more lead at the top is merely to add those bursts which are started in this new layer of lead, although it may be that these bursts do not add to the number of registered bursts because the particles are absorbed by the intervening lead *A decrease of the number of recorded bursts can only take place if less bursts are started in the lower layers of material*, which can only happen if the intensity of the electrons starting these bursts has been decreased by the superposed lead Since these electrons have energies above the critical breakdown energy they belong to the penetrating group, and their decrease will be proportional to the decrease of the *penetrating* component Thus, after the maximum, *the curve for bursts should decrease at roughly the same rate as the penetrating component*, which is contrary to observation It should be noticed that his conclusion is independent of the exact mechanism of the bursts and holds whether or not there are highly absorbable intermediate links In fact, with this mechanism the transition curve would look like curves 2, 3 and 5 for the showers which we have seen to be associated directly or indirectly with the penetrating component We conclude then that the observed shape of *the transition curve for bursts makes it highly improbable that the theory of energy loss fails for electrons in lead at energies of* 4×10^8 e-volts, and it allows us to answer the first question stated at the beginning of this section in the affirmative, i e *that electrons of sufficiently high energy do produce bursts entirely by cascade multiplication* We are still not in a position to answer the second question, for those showers and bursts which are associated with the hard component may be produced by a penetrating particle emitting a sufficiently hard light quantum, or producing a very high energy electron by collision, either of which would then produce showers according to the cascade theory

It would be possible to decide whether or not particles can produce showers in one elementary process by investigating the showers produced in sheets of material so thin that it would be impossible for sufficient multiplication of the number of particles to take place on the cascade theory in such sheets According to Bhabha and Heitler, in a thickness corresponding to 2 in the units characteristic of the material (Bhabha and Heitler 1937, eq (15)) the mean number of particles with energy greater than 10^7 e-volts produced by an electron of 2×10^{11} e-volts is 30, and since roughly an equal number of particles have an energy below 10^7 e-volts, this shower will contain on the average about 60 particles Electrons of less energy produce still smaller showers Messerschmidt (1936) has observed behind 9 cm of aluminium or 20 cm of coal in a chamber whose walls were of 0 7 cm iron, bursts of more than 200 particles with a quite comparable frequency * Since these thicknesses of absorber correspond to thicknesses less than 2, one would have to suppose that these bursts represent fluctuations in the number of particles in a shower from the mean number The chance that a shower of more than 200 particles should appear as a fluctuation when the expected average number is 60, is (Bhabha and Heilter 1937, eq (31)) of the order 10^{-45} It would thus appear as if these experiments proved without doubt that high-energy particles can in fact also create a shower of particles in one elementary process The force of these results is, however, weakened by the fact that there was a low wall at a distance of 1 5 m from the ionization chamber, and the surrounding walls were only 4 m removed from it A number of particles of a shower occurring in the walls might then quite easily hit the apparatus, thus upsetting our calculations We also do not know to what extent the particles of a shower produced, say, in the top of the chamber produce subsidiary showers in the walls by cascade multiplications A similar difficulty is met in interpreting the results of Carmichael (1936) who, in a chamber made of iron $\frac{1}{8}$ in thick, and with no heavy material above the chamber, found bursts corresponding to more than 200 particles One cannot be sure that the burst was not started in some iron girder in the roof * We nevertheless hold it for not improbable that cases do occur in which a number of particles are created in one elementary process Indeed, that such processes are to be expected on Fermi's β-ray theory has been shown by Heisenberg, although the theory has to be modified before it will give results of the right order of magnitude

---

* I wish to express my thanks to Professor Heisenberg for drawing my attention to these results in the course of an exchange of letters

C—Wilson chamber experiments

Blackett and Wilson (1937) have recently measured the energy loss of cosmic ray particles passing through lead and aluminium plates put across their Wilson chamber Their results may be summarized as follows For energies up to 2×10^8 e-volts the energy loss of cosmic-ray particles in lead is in agreement with that to be expected theoretically for electrons, confirming the earlier results of Anderson and Neddermeyer (1936) For energies greater than 2×10^8 e-volts the ratio of the experimental to the theoretical energy loss in lead decreases rapidly, reaching a value of about a quarter at energies of about 4×10^8 e-volts After this the decrease is more slow, the ratio being less than about a twentieth for energies near 4×10^9 e-volts

In aluminium the relative energy loss seems to be about one-fifth of that in lead for the energy range from 5×10^8 e-volts to 2×10^9 e-volts, although the accuracy of the aluminium measurements is admittedly not as high as that in lead

It is clear from these experiments that there is already a marked discrepancy with the theory for curvatures which correspond to an energy of 4×10^8 e-volts if the particles be electrons Since for such curvatures the ionization of a proton would be about two and a half times that on an electron, and it is claimed by Blackett and Wilson that it is possible to notice differences in the ionization of this amount, the possibility of these particles being protons may be excluded Thus these experiments compel one directly to accept either hypothesis (*a*) or (*b*)

---

* Mr Carmichael in a conversation himself drew my attention to this possibility

If the explanation of these experiments is to be found in a breakdown above some critical energy of the theory for radiative energy loss, then it is quite clear that this critical energy must depend on the atomic number of the material, for the critical energy in lead must be put between 2 and 4×10^8 e-volts, whereas we have seen that in air it cannot be below 10^{10} e-volts Indeed, the rough energy loss measurements in aluminium seem to support this view

To explain the observed energy loss in lead on hypothesis (*b*) it would be necessary to assume that at sea-level the radiation consisted of a mixture of electrons and heavy electrons, possibly having several different rest masses Particles below 4×10^8 e-volts would be mostly electrons, those above this energy mostly heavy electrons If the energy losses of all these particles be due entirely to the ionization and the ordinary Bremsstrahlung, then we would have to conclude that the majority of particles of energy round about 4×10^9 e-volts should have a rest mass of about five times the electron mass or more, since the observed radiation loss is less than about one-twentieth of that for electrons This sets a lower limit to the mass of heavy electrons with energies in the neighbourhood of 4×10^9 e-volts We cannot give an upper limit from such considerations, since the observed loss may be due to an average of the energy loss of heavy particles and a small number of electrons The fact that in the atmosphere the particles seem to show the theoretical loss for electrons up to energies as high as 10^{10} e-volts presents no difficulty, for in the upper atmosphere the soft electron component would predominate and control the shape of the absorption curve, whereas, by the time the radiation has reached sea-level, most of these electrons would have already been absorbed

But an explanation along these lines meets with the difficulty that in lighter elements, for example, aluminium, the observed energy loss would also be less than about a twentieth of the theoretical at about the same energies, whereas in fact it seems to be comparable with the theoretical energy loss Thus, if the findings of Blackett and Wilson are correct, namely, that the actual energy loss deviates from the theoretical in lead for energies above 4×10^8 e-volts whereas in aluminium no considerable deviation occurs for energies up to some much higher value, say 2×10^9 e-volts, then we are forced to the following conclusion *If the explanation of these experiments is to be sought by assuming the existence of new particles, then we must conclude that* for high energies *part of the loss is not due to the ordinary Bremsstrahlung* (which varies as Z^2), *but to some other process which allows the* *possibility of large losses besides the ionization loss, such losses varying in different substances not as Z^2 but rather as Z* One possibility for such losses is discussed in the preceding section

Recently Neddermeyer and Anderson (1937) have reported energy-loss measurements in a platinum plate put across the chamber in which they divide the tracks entering the chamber into two groups, shower particles and single tracks The tracks of the first group show the energy loss to be expected theoretically for electrons, while the particles of the second group show a much lower energy loss in the same range Since the lowest energy tracks of the second group have a curvature corresponding to an energy of about 1 4 × 10^8 e-volts if they were electrons, one may, as before, exclude the possibility of these particles being protons, since protons of this curvature would exhibit a much larger ionization Under these circumstances Neddermeyer and Anderson conclude that their experiments indicate the existence of a new particle of electronic charge and a mass intermediate between those of the electron and proton *

Lastly, Anderson and Neddermeyer have found that in certain photographs of showers, particles are seen which show an ionization definitely heavier than that of an electron On the other hand, the range of these particles is much longer than it would be for protons having the observed curvature If, therefore, no error has occurred in the estimate of the curvature or range, these photographs would supply additional evidence of a new particle of the type we are considering, as the authors themselves point out The mass of these particles would be of the order of a few hundred times the electron mass

To sum up then, we may say that while the energy-loss measurements of Blackett and Wilson would find a more simple explanation on the hypothesis (*a*) that the theory of radiation loss fails for electrons, *this hypothesis would seem to be in contradiction with other definitely established*

---

* It must be remarked however that these results are not quite in harmony with those of Blackett and Wilson Since the first group shows a normal energy loss, and the second group a lower energy loss, it follows that the energy loss averaged over particles of both groups is less than that to be expected theoretically, whereas Blackett and Wilson find that the energy loss up to 2 × 10^8 e volts is in reasonable agreement with the theory On the other hand, apart from the actual value of the energy losses for the two groups, the fact that the two groups show a markedly different energy loss seems to be clear, in accordance with the observation often made by various investigators that the particles in showers seem to show a higher energy loss and are more absorbable than those not in showers In passing we may remark that the suggestion that has been made, that the shower particles have a mass smaller than that of the electron, may be rejected on the ground that a gamma ray would have a greater chance of producing a pair of such particles than an electron pair, and the threshold frequency for pair creation would also be lower than 10^6 e volts

phenomena connected with the latitude effect and the transition curves for bursts In these circumstances we seem to be compelled to accept the hypothesis (*b*) of the existence of new particles

2—ENERGY LOSS

Free collision and ionization

We now proceed to investigate as far as possible the behaviour of a particle, of charge e equal to the electronic charge, and of some arbitrary mass M between those of the electron and proton We shall assume that the particle obeys the Dirac equation, and that its interaction with other charged particles is that given by quantum electrodynamics for the interaction of point charges The correctness of our results will then be limited by two possibilities First, the particle may have a direct interaction with other particles like itself or even with electrons other than that operating through the electromagnetic field, as is indeed the case for protons We, however, regard it as unlikely that such interactions, if they exist, would affect the ionization loss appreciably, since the important contribution to the loss comes from processes which take place at large distances, and direct interactions between particles are usually short-range forces Secondly, the interaction of the particle with the radiation field may differ from that due to a point charge owing to the particle possessing something corresponding in the classical picture to its charge being spread over a region of finite extension This will again not affect the ionization loss appreciably, but, as we shall see when we come to consider radiation loss, it may impose restrictions on the validity of the radiation formulae which are more stringent than those for electrons since the "Compton wave-length" $\hbar/Mc$ of such a particle is less than the Compton wave-length $\hbar/mc$ of electrons

We shall first consider the cross-sections for the collision of a "heavy electron" with an ordinary electron at rest in the material The general expression for this cross-section taking retardation into account has been given by Møller (1932), and it can be evaluated exactly as has been done in a previous paper (Bhabha 1936) We give only the result here The differential effective cross-section $Q\,dq$ for the production by a "heavy electron" of total energy E_0 of a secondary electron of kinetic energy W lying in the interval corresponding to dq, where

$$q \equiv W/(E_0 - Mc^2)$$

and

$$\gamma \equiv E_0/Mc^2,$$

is

$$Q(E_0, W)\,dW \equiv Q(q)\,dq = 2\pi r_0^2 \frac{m}{M(\gamma-1)}\left[\frac{\gamma^2}{\gamma^2-1} - \frac{q}{q_m} + \frac{1}{2}\left(\frac{\gamma-1}{\gamma+1}\right)q^2\right]\frac{dq}{q^2} \qquad (2\ 1)$$

Here m is the mass of the electron, $r_0 \equiv e^2/mc^2$, and

$$\frac{W_m}{E_0 - Mc^2} \equiv q_m = \frac{2mM(\gamma+1)}{m^2 + M^2 + 2mM\gamma} \qquad (2\ 2)$$

q may take on all values consistent with the conservation of energy and momentum, namely, from 0 to q_m Then $W_m \equiv q_m(E_0 - Mc^2)$ is the maximum energy which can be transferred to an initially stationary electron in a free collision, and it plays a certain role in our later calculations It is clear that q_m tends to unity for sufficiently large γ however great M may be

We must emphasize that just the cross-sections for hard collisions given by expression (2 1) might be considerably altered by the existence of close-range forces In the absence of any knowledge about such forces and a relativistic formulation of them, we cannot estimate the magnitude of this correction We would mention in passing that the spin of the particle also plays an important role for just these hard collisions, and since the spin of the proton, for example, is not described completely by the Dirac equation, it is of interest to estimate the order of this correction If we had described the heavy particles by the Klein-Gordon equation instead of the Dirac equation, thus ascribing no spin to them, the last term in square brackets in (2 1) would have been replaced by

$$\frac{1}{\gamma+1}\frac{m}{M}q$$

The difference due to the spin therefore bears a ratio to the total cross-section of the order $(\gamma^2 - 1)\,m^2/M^2$, which for protons is small compared with unity except for energies of the order of 10^{12} e-volts

The total ionization loss per centimetre is given by the formula of Bloch

$$-\left(\frac{dE_0}{dx}\right)_{\text{coll}} = 2\pi\sigma Z r_0^2 \frac{mc^2}{\beta^2}\left[\log\frac{mc^2\beta^2 W_m}{(1-\beta^2)\,I^2 Z^2} + (1-\beta^2)\right], \qquad (2\ 3)$$

where σ is the number of atoms per cubic centimetre of the substance, Z the atomic number, $c\beta$ the velocity of the heavy electron given by $\beta = \sqrt{(1 - 1/\gamma^2)}$ and IZ the mean ionization potential of the atom, where we put, with Bloch, $I = 13\ 5$ e-volts W_m has been defined above Our results are given in Tables II and III

Radiation loss and creation of pairs

We mentioned in the preceding section that the quantum-mechanical formulae for the radiation loss of heavy electrons are not on as sound a footing as for ordinary electrons To see this we consider the derivation of the formulae by the method of Weizsäcker and Williams We consider the whole process in a Lorentz system in which the radiating particle is at rest In this system the nucleus moves along a straight line with extremely high velocity, and its field may then be considered as a superposition of quanta of different frequencies The radiation process corresponds in this system to the scattering of one of these quanta by the stationary particle Weizsäcker has shown that the important contribution to the radiation comes from the scattering of all quanta whose energy is equal to or less than the rest energy of the particle In other words, the validity of the radiation formulae depends on the validity of the Klein-Nishina formula for wave-lengths as small as the "Compton wave-length" of the particle Now in the case of electrons, the Compton wave-length $\hbar/mc$ is large compared to the classical radius of the electron e^2/mc^2, so that the validity of the Klein-Nishina formula is not in question The classical radius e^2/Mc^2 of a heavy electron would also be small compared to its Compton wave-length $\hbar/Mc$, so that in this case, too, no difficulty would arise We, however, think that it may be a property of charge in general that it may not be possible to localize it in a region smaller than e^2/mc^2, whatever the mass of the particle with which it is associated In this case the Klein-Nishina formula may no longer be valid for quanta of energy equal to the rest energy of "heavy electrons" if

$$\frac{\hbar}{Mc} \lesssim \frac{e^2}{mc^2},$$

i e if

$$\frac{m}{M} \lesssim \frac{e^2}{\hbar c} \tag{2 4}$$

Thus we must be prepared to find that for particles of mass greater than 137m the radiation formulae may not be valid

It is quite easy to calculate the radiation loss for heavy electrons, taking screening into account to the same degree of accuracy as in the original calculations of Bethe and Heitler The calculations follow those of Bethe (1934) closely It can be shown exactly, as has been done there, that the differential effective cross-section for the emission of a quantum k in the energy interval dk by a heavy electron of energy E_0 is

$$\begin{aligned}\phi(k)\,dk = \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2\frac{4dk}{E_0^2k}\Bigg[(E_0^2+E^2)\left\{\int_\delta^{Mc^2}(1-F)^2(q-\delta)^2\frac{dq}{q^3}+1\right\} \\ -\tfrac{2}{3}E_0E\left\{\int_\delta^{Mc^2}(1-F)^2\left(q^3-6\delta^2q\log\frac{q}{\delta}+3\delta^2q-4\delta^3\right)\frac{dq}{q^4}+\frac{5}{6}\right\}\Bigg],\end{aligned} \tag{2 5}$$

where $\delta = \frac{(Mc^2)^2 k}{2E_0 E}$ We shall assume throughout that

$$E_0,\ E,\ k \gg Mc^2,$$

for it is only then that screening is important, so that

$$\delta \ll Mc^2$$

The form factor F which represents the effect of screening is given by

$$F\left(\frac{q\mu Z^{-\frac{1}{3}}}{c\hbar}\right) \equiv \frac{1}{4\pi}\int e^{\frac{i}{c\hbar}\mu Z^{-\frac{1}{3}}(q,x)} \frac{\phi^{\frac{3}{2}}(r)}{r^{\frac{3}{2}}} dx \tag{2 6}$$

and is a function of $q\mu Z^{-\frac{1}{3}}/c\hbar$ only The integration in (2 6) extends over the whole of space ϕ is the function tabulated by Fermi, and determines the density of electrons at a distance r from the nucleus in the units of length defined by μ, where

$$\mu = \left(\frac{3\pi}{8\sqrt{2}}\right)^{\frac{2}{3}} \frac{\hbar^2}{me^2} = 0\ 876 \frac{\hbar^2}{me^2} \tag{2 7}$$

Now ϕ is only considerable when r is less than or of the order unity, so that $F \ll 1$ when

$$q \gg \frac{c\hbar}{\mu} Z^{\frac{1}{3}} \sim \frac{e^2}{\hbar c} Z^{\frac{1}{3}} mc^2 \tag{2 8}$$

Screening is therefore only effective when δ, the minimum value of q, satisfies

$$\delta \lesssim \frac{e^2}{\hbar c} Z^{\frac{1}{3}} mc^2 \ll mc^2 \tag{2 8a}$$

Hence if the inequality (2 8a) is not satisfied, i e if

$$\frac{Mc^2 k}{2E_0 E} \frac{M}{m} \gg Z^{\frac{1}{3}} \frac{e^2}{\hbar c}, \tag{2 8b}$$

we may neglect F in (2 5) altogether, and the integrations can then be carried out exactly We now consider the first q integral in (2 5) when (2 8a) is satisfied We may write it

$$\int_{\delta}^{mc^2} (1-F)^2 (q-\delta)^2 \frac{dq}{q^3} + \int_{mc^2}^{Mc^2} (1-F)^2 (q-\delta)^2 \frac{dq}{q^3} \tag{2 9}$$

In virtue of (2 8) and (2 8a) both F and δ may be neglected in the second integral in (2 9) and it becomes $\log M/m$ The first integral has been evaluated numerically by Bethe and may be written in the form

$$\tfrac{1}{4}\phi_1(\xi) + \log Z^{-\frac{1}{3}},$$

where

$$\xi = \frac{200}{mc^2} Z^{-\frac{1}{3}}\delta = \frac{100}{Z^{\frac{1}{3}}} \frac{Mc^2k}{E_0 E} \frac{M}{m} \tag{2 10}$$

(2 9) therefore reduces to

$$\tfrac{1}{4}\phi_1(\xi) + \log \frac{M}{m} Z^{-\frac{1}{3}}$$

The second q integral in (2 5) can be treated similarly We get finally

$$\phi(k)\,dk = \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \frac{dk}{E_0^2 k}\left[(E_0^2 + E^2)\left\{\phi_1(\xi) + 4\log\frac{M}{m} Z^{-\frac{1}{3}}\right\} - \tfrac{2}{3}E_0 E\left\{\phi_2(\xi) + 4\log\frac{M}{m} Z^{-\frac{1}{3}}\right\}\right] \tag{2 11}$$

Here ξ is defined by (2 10), and the functions ϕ_1 and ϕ_2 have been given by Bethe and Heitler (1934, fig 1) Moreover, as Bethe (1934, eq (63)) has shown, for the case of complete screening $\delta = 0$

$$\phi_1(0) = \phi_2(0) + \tfrac{2}{3} = 4\log 183,$$

so that (11) reduces to

$$\phi(k)\,dk = 4\frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \frac{dk}{E_0^2 k}\left\{(E_0^2 + E^2 - \tfrac{2}{3}E_0 E)\log 183\frac{M}{m} Z^{-\frac{1}{3}} + \tfrac{1}{9}E_0 E\right\} \tag{2 12}$$

For low energies when the inequality (2 8*b*) is satisfied, i e when $\xi \gg 1$,

$$\phi_1(\xi) = \phi_2(\xi) = 4\left(\log\frac{200}{\xi} - \frac{1}{2}\right),$$

so that (2 11) reduces to

$$\phi(k)\,dk = 4\frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \frac{dk}{E_0^2 k}\left\{(E_0^2 + E^2 - \tfrac{2}{3}E_0 E)\left(\log\frac{2E_0 E}{kMc^2} - \frac{1}{2}\right)\right\}, \tag{2 13}$$

as indeed a direct integration of (2 5) neglecting the factors F would have given

The average energy loss by radiation per centimetre is given by

$$-\left(\frac{dE_0}{dx}\right)_{\text{rad}} = \sigma\int_0^{E_0} k\phi(k)\,dk, \tag{2 14}$$

where σ is the number of atoms per c c and $\phi(k)$ is given by (2 11) Introducing the variable $\epsilon = k/E_0$, it may be written in the form

$$\sigma \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 E_0 \int_0^1 d\epsilon \left[\{1+(1-\epsilon)^2\}\phi_1\left(\frac{50}{\xi'}\frac{\epsilon}{1-\epsilon}\right) - \tfrac{2}{3}(1-\epsilon)\phi_2\left(\frac{50}{\xi'}\frac{\epsilon}{1-\epsilon}\right) + 4\log\frac{M}{m} Z^{-\frac{1}{3}}\{\tfrac{4}{3}(1-\epsilon)+\epsilon^2\}\right], \quad (2\ 14a)$$

where

$$\xi' = Z^{\frac{1}{3}}\left(\frac{m}{M}\right)^2\left(\frac{E_0}{2mc^2}\right) \quad (2\ 15)$$

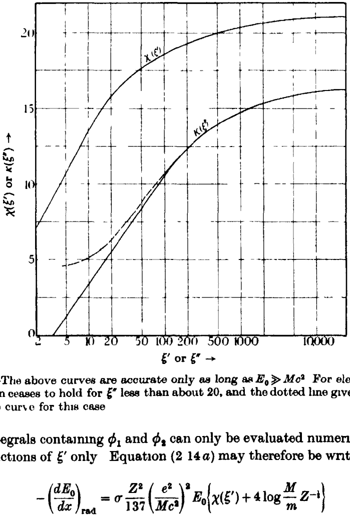


FIG 2—The above curves are accurate only as long as $E_0 \gg Mc^2$ For electrons this condition ceases to hold for ξ'' less than about 20, and the dotted line gives the more accurate curve for this case

The integrals containing ϕ_1 and ϕ_2 can only be evaluated numerically, and are functions of ξ' only Equation (2 14a) may therefore be written in the form

$$-\left(\frac{dE_0}{dx}\right)_{\text{rad}} = \sigma\frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 E_0\left\{\chi(\xi') + 4\log\frac{M}{m} Z^{-\frac{1}{3}}\right\} \quad (2\ 16)$$

We have plotted χ as a function of ξ' in fig 2, where ξ' defined by (2 15) is the energy measured in millions of volts multiplied by the factor $Z^{\frac{1}{3}}(m/M)^2$ Since for the case of electrons, $M = m$, the expression in curly brackets must reduce to ϕ_{rad} already calculated by Bethe and Heitler (1934, eq (48)), the values of χ can be deduced from those of ϕ_{rad} For $\xi' \ll 137$ (negligible screening)

$$\chi(\xi') \rightarrow 4 \log 4\xi - \tfrac{4}{3},$$

and for $\xi' \gg 137$ (complete screening)

$$\chi(\xi') \rightarrow 4 \log 183 + \tfrac{2}{9},$$

so that in the first case

$$-\left(\frac{dE_0}{dx}\right)_{\text{rad}} = \sigma \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 E_0\left\{4 \log \frac{2E_0}{Mc^2} - \frac{4}{3}\right\}, \tag{2 16a}$$

and in the second case

$$-\left(\frac{dE_0}{dx}\right)_{\text{rad}} = \sigma \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 E_0\left\{4 \log 183 \frac{M}{m} Z^{-\frac{1}{3}} + \frac{2}{9}\right\} \tag{2 16b}$$

We therefore see that the larger the mass of the particle, the later screening becomes effective and the less its effect

In Tables II and III we give the collision and radiation losses in lead and water of particles of masses 10 and 100 times the electron mass for various energies The total energy loss is shown in fig 3

TABLE II—ENERGY LOSS IN LEAD

Millions of electron volts per centimetre

$E_0 - Mc^2$		10^6	10^7	10^8	10^9	10^{10}	10^{11}	10^{12}
Electron	Coll	11 4	13 9	18 6	23 4	28 2	33 0	37 7
	Rad	–	14 4	177	1900	19,400	$1\ 94 \times 10^5$	$1\ 94 \times 10^6$
$Mc^2 = 5 \times 10^6$ e volts	Coll	26 4	12 9	16 0	20 8	25 4	30 3	35 0
$\approx 10mc^2$	Rad	—	—	1 72	26 4	296	3080	31,100
$Mc^2 = 5 \times 10^7$ e volts	Coll	132	26 8	13 2	16 9	22 2	27 1	31 8
$\approx 100mc^2$	Rad	—	—	—	0 17	2 9	38 1	413
Proton	Coll	(537)	203	40 3	13 9	15 7	21 7	27 4
$M = 1840m$	Rad	—	—	—	—	—	0 07	1 01

Using (2 3) and (2 16) we find that the ratio of the radiation to the collision loss is given by

$$\frac{-(dE_0/dx)_{\text{rad}}}{-(dE_0/dx)_{\text{coll}}} = \frac{ZE_0}{1300mc^2}\left(\frac{m}{M}\right)^2 \tag{2 17}$$

TABLE III—ENERGY LOSS IN WATER

Millions of electron volts per centimetre

$E_0 - Mc^2$		10^6	10^7	10^8	10^9	10^{10}	10^{11}	10^{12}
Electron	Coll	1 93	2 15	2 72	3 29	3 94	4 53	5 10
	Rad	—	0 16	2 07	22 5	233	$2\ 33 \times 10^3$	$2\ 33 \times 10^4$
$Mc^2 = 5 \times 10^6$ e volts	Coll	4 76	2 10	2 44	3 02	3 61	4 19	4 77
$\approx 10mc^2$	Rad	—	—	0 02	0 28	3 26	34 5	350
$Mc^2 = 5 \times 10^7$ e volts	Coll	28 1	4 82	2 14	2 54	3 20	3 80	4 39
$\approx 100mc^2$	Rad	—	—	—	—	0 03	0 40	4 45
Proton	Coll	(284)	47 1	7 48	2 32	2 40	3 13	3 84
$M = 1840m$		Radiation loss negligible						

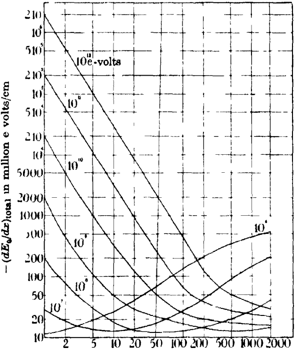


FIG 3—Energy loss in lead

By putting (2 17) equal to unity, we may define a critical energy at which the rates of radiation and collision losses are equal, namely

$$E_M = \frac{1300}{Z}\left(\frac{M}{m}\right)^2 mc^2 \qquad (2\ 18)$$

For energies greater than this the loss is predominantly due to radiation, for smaller energies to collision This energy is correctly given by (2 18) to within 20 % More accurately, the numerical constant in (2 18) should be 1600 for $M = m$, 1300 for $M = 10m$ and 1100 for $M = 100m$

The total rate of energy loss of the particle may be roughly written in the form

$$-\left(\frac{dE_0}{dx}\right)_{\text{total}} = \frac{Z^2}{137} r_0^2 \sigma \left(\frac{m}{M}\right)^2 C_M [E_0 + E_m], \qquad (2\ 19a)$$

where C_M is a numerical constant,

$$C_M = 23 \text{ for } M = 10m, \quad \text{and} \quad C_M = 30 \text{ for } M = 100m$$

The mean range R is then roughly given by

$$R = \frac{137}{Z r_0^2 \sigma \left(\frac{m}{M}\right)^2 C_M} \log\left(1 + \frac{E_0 - Mc^2}{E_M + Mc^2}\right), \qquad (2\ 19b)$$

which is accurate to within 30 % except for kinetic energies small compared to the rest mass of the particle (2 19b) would give a better fit with the numerically calculated ranges if we chose slightly different values for E_M, putting the numerical constant in (2 18) equal to 1100 for $M = 10m$ and 920 for $M = 100m$ respectively

The cross-sections for the creation of pairs by γ-rays are obtained from the above formulae if we change the sign of the energy E_0 of the initial state, writing in its place $-E_+$, E_+ being the energy of the positron, and replacing the factor $dk/E_0^2 k$ by dE/k^3 We thus get for the cross-section for the creation by a quantum $k = h\nu$ of a pair, the electron of which has an energy in dE, the expression

$$\phi(E)\,dE = \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \frac{dE}{(h\nu)^3}\left[(E^2 + E_+^2)\left\{\phi_1(\xi) + 4\log\frac{M}{m}Z^{-\frac{1}{3}}\right\}\right.$$
$$\left. + \tfrac{2}{3} E E_+ \left\{\phi_2(\xi) + 4\log\frac{M}{m}Z^{-\frac{1}{3}}\right\}\right] \qquad (2\ 20)$$

The total cross-section for pair creation is given by

$$\phi_{\text{pair}} = \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \left[\kappa(\xi'') + \frac{28}{9}\log\frac{M}{m}Z^{-\frac{1}{3}}\right] \qquad (2\ 21)$$

The function $\kappa(\xi'')$ is plotted in fig 2, where

$$\xi'' = Z^{\frac{1}{3}}\left(\frac{m}{M}\right)^2\left(\frac{h\nu}{2mc^2}\right) \qquad (2\ 22)$$

When $\xi'' \ll 137$ (negligible screening)

$$\kappa(\xi'') \to \tfrac{28}{9} \log 4\xi'' - \tfrac{218}{27},$$

and for complete screening $\xi'' \gg 137$

$$\kappa(\xi'') \to \tfrac{28}{9} \log 183 - \tfrac{2}{27},$$

so that (2 21) becomes

$$\phi_{\text{pair}} = \frac{Z^2}{137}\left(\frac{e^2}{Mc^2}\right)^2 \begin{cases} \left\{\dfrac{28}{9} \log \dfrac{2h\nu}{Mc^2} - \dfrac{218}{27}\right\}, \\ \left\{\dfrac{28}{9} \log 183 \dfrac{M}{m} Z^{-\frac{1}{3}} - \dfrac{2}{27}\right\}, \end{cases} \tag{2 23}$$

in the two cases respectively

3—THE SECONDARIES ACCOMPANYING HEAVY PARTICLES

A heavy particle in its passage through matter will directly or indirectly produce secondary electrons which on account of their lower energy and lighter mass will behave like a soft component accompanying the penetrating heavy particle It is therefore of interest to calculate the number of electrons or positrons of energy greater than some value E in equilibrium with a homogeneous beam of heavy particles of mass M and energy E_0

The number of secondaries produced by direct hard encounters is given by (2 1) But this is not the quantity, which is of direct interest in cosmic-ray experiments These direct secondaries will produce further positive and negative electrons by cascade multiplication in the subsequent layers of material In addition to this, the heavy particle may emit a quantum of radiation which will also produce electrons by cascade multiplication The average number of electrons which emerge from the bottom of some layer of substance through which the heavy particle has passed is therefore different from that given by a simple application of (2 1) We now proceed to estimate this number

We will first calculate the average number of positrons and electrons which accompany the passage of a heavy particle of energy E_0, and which are due only to the direct production of fast secondaries by collision as given by (2 1) and the ensuing cascade processes We will suppose that the layer already traversed by the heavy particle is so thick that a cascade process started at the beginning of the layer does not reach the point we are considering Such a layer may be described as "infinitely thick" A layer satisfying this condition has a thickness of about 30 in the units λ_0 characteristic of the material in the cascade theory λ_0 is given by (Bhabha and Heitler 1937 eq (15))

$$\lambda_0 = \frac{137}{aZ^2 r_0^2 \sigma}, \qquad (3\ 1)$$

with a equal to 20 for lead and 23 for air or water, so that an infinitely thick layer is about 10 m of water or 12 cm of lead Under these circumstances it can be shown (Arley and Bhabha 1937) that the number $N_{s>}$ of electrons (+ and −) with energy greater than E is given by

$$\begin{aligned} N_{s>} = \frac{2\pi}{aZ}\frac{137\,mc^2}{E}\Bigg[&\frac{\gamma^2}{\gamma^2-1}\left\{1-\frac{1}{r}(\log r+1)+2\alpha\left[\frac{1}{\beta-1}(r^{\beta-1}-1)+\frac{1}{r}-1\right]\right\} \\ &-\frac{(\log r)^2}{2r}-\frac{2\alpha}{r}\left\{\frac{1}{\beta}(r^{\beta}-1)-\log r\right\} \\ &+\frac{1}{2}\frac{\gamma-1}{\gamma+1}q_m^2\left\{\frac{\log r-1}{r}+\frac{1}{r^2}+\frac{2\alpha}{r}\left(\frac{1}{\beta+1}[r^{\beta+1}-1]-r+1\right)\right\}\Bigg], \qquad (3\ 2) \end{aligned}$$

where $r = W_m/E$, and $\alpha = 0\ 224$ and $\beta = 1\ 029$ are numerical constants W_m is defined in (2 2) This is valid for $E > E_c$, where E_c is the "critical energy" for the substance, being the energy at which the ionization loss of an electron is equal to the radiation loss This limit is (Bhabha and Heitler 1937, § 1)

$$10^7 \text{ e-volts in lead, } 1\ 5 \times 10^8 \text{ e-volts in air or water} \qquad (3\ 3)$$

(3 2) is accurate to the same degree as the calculations of the cascade theory, i e to within about 30%, except when $W_m \sim E_c$

To estimate roughly the number of electrons below E_c we use the result of the cascade theory discussed in § 4, namely, that in a shower the number of particles with energy below E_c is very roughly equal to the number above E_c We may thus deduce the following expression for the number $N_{s<}$ of positrons and electrons with energy below E_c accompanying the heavy particle,

$$\begin{aligned} N_{s<} = \frac{2\pi}{aZ}\frac{137\,mc^2}{E_c}\Bigg[&\left\{1-\frac{1}{r}+\frac{2\alpha\beta}{\beta-1}(r^{\beta-1}-1)\right\}\frac{\gamma^2}{\gamma^2-1}-\frac{1}{r}\log r-\frac{2\alpha}{r}(r^{\beta}-1) \\ &+\frac{1}{2}\frac{\gamma-1}{\gamma+1}\frac{q_m^2}{r^2}\left\{r-1+\frac{2\alpha\beta}{\beta+1}(r^{\beta+1}-1)\right\}\Bigg] \qquad (3\ 4) \end{aligned}$$

We give the values of (3 2) and (3 4) in Table IV

We see that for extremely high energies (10^{12} e-volts) the number of secondaries *due to this process alone* is practically independent of the mass of the particle, being about 33 % of the hard component in lead For lower energies it depends on the mass to a greater extent For particles with a kinetic energy of 10^{10} e-volts passing through lead the secondary electron component would constitute about 20 % for particles of ten to a hundred times the electron mass, whereas it is about 7 % for protons In water the secondary electron component varies from about 7 % for particles with kinetic energy of 10^{10} e-volts to about 15 % for particles of 10^{12} e-volts According to Auger and his co-workers, the soft component in water which is in equilibrium with the penetrating component is of the order of 5 % or less,* which allows one to deduce from Table IV, that the mean energy of the penetrating particles must be of the order of 10^{10} e-volts *or less*, as is in fact otherwise known It must be emphasized that in estimating the equilibrium intensity, the number of particles appearing in showers must be *included* in calculating the average We see that the amount of the secondary electron component in equilibrium with the penetrating component as given by the theory is in qualitative agreement with experiment, although both the theoretical and experimental results are too inaccurate to allow of a quantitative comparison

TABLE IV—THE MEAN NUMBER PER CENT OF POSITRONS AND ELECTRONS WITH ENERGY GREATER THAN E_c ($N_{s>}$) AND ENERGY LESS THAN E_c ($N_{s<}$) ACCOMPANYING THE HEAVY PARTICLE

The addition of secondaries from radiative losses would roughly double the figures for $M = 10m$ (figures in brackets)

		Lead			Air or water		
$E_0 - Mc^2$ in mV		10^8	10^{10}	10^{12}	10^8	10^{10}	10^{12}
$Mc^2 = 5 \times 10^6$ e volts	$N_{s>}$	1 8	9 4	16	—	3 0	7 2
$\approx 10mc^2$	$N_{s<}$	3 0	11 0	18	—	3 9	8 1
	Total	4 8	20 4	34	—	6 9	15 3
		(10)	(41)	(70)	—	(14)	(31)
$Mc^2 = 5\ 10^7$ e volts	$N_{s>}$	—	9 0	16	—	2 9	7 2
$\approx 100mc^2$	$N_{s<}$	—	10 0	18	—	3 9	8 1
	Total	—	19 0	34	—	6 8	15 3
Proton	$N_{s>}$	—	2 5	15 1	—	—	6 5
$M = 1840m$	$N_{s<}$	—	4 1	16 7	—	—	7 4
	Total	—	6 6	31 8	—	—	13 9

We must add to the figures of Table IV the number of electrons which result from cascade processes caused by quanta emitted by the heavy particle It will appear in § 4, eq (4 1), that the ratio of the relative probabilities of the emission of a quantum and of the production of an electron by collision both with energy greater than E_c is of the order unity for $M = 10m$ and is small compared to one for $M \gg 10m$ A quantum, moreover, produces roughly the same number of particles by cascade multiplication as an electron, so that in Table IV the figures for $M = 100m$ and for protons will not be appreciably altered, *whereas the figures for $M = 10m$ will be somewhat more than doubled by the addition of this process* A comparison with the measurements of Auger and others therefore allows one to conclude that if the hypothesis of new particles is right *the majority of the penetrating particles must have masses nearer to a hundred times the electron mass rather than ten times the same*

---

* This equilibrium intensity must be taken from measurements under water, for as we have pointed out in § 1, the two components are not in equilibrium above sea level

4—NUMBER AND SIZE OF SHOWERS

We now proceed to investigate the production of showers by heavy particles There are three distinct ways in which a heavy particle may cause a shower (1) It may produce a very fast secondary electron by direct collision which in its turn may produce a shower by cascade multiplication if it has sufficient energy (2) It may emit a quantum of radiation of sufficient energy to produce a shower by cascade multiplication * (3) It may produce a shower directly by a multiple process The second process may itself occur in two distinct ways The emitted quantum may be just the ordinary Bremsstrahlung, which is emitted during a change in the motion of the particle as a whole in a given external field The probability of this process has been calculated in § 2 But in addition to this we must be prepared to find that under certain circumstances a particle may change its rest mass, the difference in energy being liberated as a quantum of radiation Present quantum mechanics of course does not enable one to calculate the probability of such a process We shall not concern ourselves here with processes of the type (3) Heisenberg has shown how they may result from a modification of Fermi's β-ray theory

As has been shown by Bhabha and Heitler, cascade multiplication of the number of particles is only effective provided the particles or quanta have energies above the critical energy E_c for the material under consideration, although a certain amount of multiplication may also take place below this energy. We will begin by estimating the relative probability of a shower being produced by processes (1) and (2), which amounts to estimating the ratio of the chance of a heavy particle emitting a quantum of energy greater than E to the chance of its producing by collision an electron of energy greater than the same amount. The cross-sections for these two processes are given by formulae (2.11) and (2.1) respectively. It can be shown by an easy calculation that this ratio is

---

* Since this paper was sent to press, a note has appeared by Landau and Rumer (1937) in which they estimate the probability of showers due to process (2) Our calculations below show that for $M > 10m$ the process (1) becomes much more important than the process (2), so that for $M \approx 100m$, we get shower probabilities which are some hundred times larger than those of Landau and Rumer

$$\frac{\int_{E}^{E_0}\phi(k)\,dk}{\int_{E/E_0}^{q_m}Q(q)\,dq}=\frac{2}{\pi}\,\frac{Z}{137}\left(\frac{m}{M}\right)^2\left(\frac{E}{mc^2}\right)\log q\,.\,\log X \tag{4.1}$$

to a fair approximation provided $\gamma^2 \gg 1$ and $q \equiv \dfrac{E}{E_0-Mc^2} \ll q_m$, i.e. provided $E \ll W_m$, the maximum energy which can be communicated to an electron in a free collision. Here

$$X \equiv \begin{cases} 183Z^{-\frac{1}{3}} & \text{when } \dfrac{2E_0}{Mc^2}\dfrac{1-q}{q} > 183Z^{-\frac{1}{3}}, \\ \dfrac{2E_0}{Mc^2}\,\dfrac{1-q}{q} & \text{when } \dfrac{2E_0}{Mc^2}\dfrac{1-q}{q} < 183Z^{-\frac{1}{3}} \end{cases} \tag{4.1a}$$

Even when E is not very small compared to W_m, (4.1) still gives the correct order of magnitude. Since cascade multiplication of the number of particles is only important provided the electrons or quanta have energies above the critical energy E_c for the material under consideration, although a certain amount of multiplication may also take place below this energy, the above formula is of interest when $E \geqslant E_c$. Putting $E = E_c \approx 1600mc^2/Z$ in (4.1) we find that the ratio (4.1) is less than unity if

$$\frac{M}{m} > 2.73\sqrt{(\log q_c\,.\,\log X)}, \tag{4.2}$$

where $q_c = E_c/(E_0-Mc^2)$. The right-hand side of (4.2) is practically independent of the atomic number of the material and is roughly equal to 9 for $E_0 \sim 10^9$ e-volts and 14 for $E_0 \sim 10^{11}$ e-volts. Therefore if the mass of the heavy particle $M \gg 10m$ the effect of radiative processes is small compared with the effect of direct collisions in producing showers and electronic secondaries. For $M \lesssim 10m$ the radiative process is important, and may even predominate over the other.

We will now consider the following question. When a particle of mass M emerges after its passage through a plate of some substance of atomic number

Z and thickness l in the units (3 1) with energy E_0, what is the probability of its being accompanied by N electrons and positrons? We will first consider only the process (1), namely, the production of a fast secondary electron by collision and the subsequent cascade process The chance of an electron of energy E in the interval dE being provided in the thin layer dl', l' being measured from the lower surface of the material from which the particle emerges is just

$$\lambda_0 \sigma Z Q(E_0'(l'),\, E')\, dE'\, dl',$$

using (2 1) As a result of this electron, $F(l', y')$ electrons and positrons will emerge *on the average* by cascade multiplication from the lower surface of the material ($l' = 0$) with an energy greater than E, where $y' = \log E'/E$ The function F is given by

$$F(l', y') \equiv W(l', y') + 2f_-(l', y') \tag{4 3}$$

Here $W(l', y')$ is the incomplete γ-function and $f_-(l', y')$ is the function calculated by Bhabha and Heitler (1937, eqq (22) and (23) and Table I) The probability of N particles appearing instead of F by a fluctuation is then (Bhabha and Heitler, eq (31))

$$e^{-F}\frac{F^N}{N!} \tag{4 4}$$

The total probability of the heavy particle appearing accompanied by N positrons and electrons of energy greater than E is therefore

$$P(N) = \lambda_0 \sigma Z \int_0^l dl' \int_E dE'\, e^{-F(l',y')} \frac{F(l',y')^N}{N!} Q(E_0'(l'),\, E') \tag{4 5}$$

We may neglect the small variation in the energy of the heavy particle in thicknesses of the plate in which F is considerable and put $E_0'(l') = E_0$ We may then write

$$P(N) = \lambda_0 \sigma Z \int_E^{E_m} dE'\, Q(E_0,\, E')\, J_N(y'), \tag{4 6}$$

where
$$J_N(y) = \int_0^l dl'\, e^{-F(l',y)} \frac{F(l',y)^N}{N!} \tag{4 7}$$

As before, we are interested mainly in large thickness of material, so that we may put $l = \infty$ in (4 7), since for $l > 30$ no significant contribution comes to the integral It does not seem possible to proceed further analytically

We have worked out numerically the values of J for several values of N and y and they are given in Table V

TABLE V—VALUES OF $J_N(y)$

N \ y	0	1	3	5	7	10
1	0	(1 2)	3 03	2 40	2 60	3 05
2	0	(0 43)	2 0	1 78	1 40	1 85
5	0	0	0 48	1 18	0 65	0 70
10	0	0	0	0 95	0 43	0 38
50	0	0	0	0	0 23	0 11

The accuracy of the figures is not high, since the function F is itself only known to within about 30 % It is easy to see the general form of the dependence of J on N and y The function $F(l, y)$ for a given y has the value of 1 for $l = 0$, increases rapidly to a maximum for an l between 4 and 12, and then falls away somewhat more slowly to zero The actual value at the maximum depends on the value of y and increases rapidly with increasing y being determined by the equation

$$F_m = 0\ 062e^{0 \cdot 93y} \tag{4 8}$$

(Bhabha and Heitler, eq (28)) for y larger than about 3 Moreover, the expression (4 4) has a maximum when $F = N$ and is small when F differs considerably from this value Thus, for a given N, if y be so small that $F_m \ll N$ the integrand of (4 7) will always remain small, and hence the value of J will be small As y increases the value of F_m increases rapidly, but the value of J will remain small until a y is reached such that $F_m \simeq N$ For all values of y equal to and less than this, the integrand in (4 7) has but one maximum, namely, at the point at which $F = F_m$ As y increases still further, however, F_m becomes larger than N, and the integrand shows two maxima For still larger y such that $F_m \gg N$ the contribution to the integral (4 7) comes from two quite separate regions, one for small l' when F is increasing and is in the neighbourhood of N, and the other for large l' when F is decreasing and is also in the neighbourhood of N Thus as y increases beyond the point for which $F_m \gtrsim N$ the positions of the two maxima in the integrand of (4 7) change, but the value of the integral does not alter appreciably Due to the form (4 4) in which F occurs, the contribution to the integral (4 7) of two regions of l' in which $F \sim N$ is not the same when the two regions overlap as when they are separate The value of J is largest when the two regions overlap for then the integrand of (4 7) is considerable for the largest domain of l' Hence it follows that the largest value of J occurs for some y such that $F_m \sim CN$, where C is a numerical constant somewhat greater than unity Table V clearly displays this behaviour of J For a given N and y small, J is negligible It rises rapidly to its maximum for values of y near some value y_N, say, depending on N, after which it decreases a little, but nevertheless remains of the same order of magnitude Further, for large N all values of J are smaller than for small N since then the maximum of the expression (4 4) is itself smaller

For a rough estimate, therefore, we may put $J = 0$ in (4 6) for $y' < y_N$, i e /or $E' < E_N$, and equal to a constant value $J_N(y_N)$ for $y' \geqslant y_N$ or $E' \geqslant E_N$, where y_N and E_N are the values of y and E respectively at which the maximum of J occurs We may determine them approximately by the condition that the maximum number of particles F_m which can be created for this y as given by (4 8) shall be equal to CN, C being a numerical constant somewhat larger than unity This gives

$$\left(\frac{E_N}{E}\right)^{0\ 93} \equiv e^{0\ 93 y_N} = \frac{CN}{0\ 062},$$

or

$$E_N = 1\ 67(CN)^{1\ 075}\,E \qquad (4\ 9)$$

Using the expression (2 1) for $Q(E_0, E')$ and (3 1) for λ_0 we get

$$P(N) \approx \frac{2\pi\ 137}{aZ}\ \frac{mc^2}{1\ 67(CN)^{1\ 075}\,E}\,J_N(y_N)$$
$$\left[\frac{\gamma^2}{\gamma^2-1}(1-u)+u\log u+\frac{1}{2}\,\frac{\gamma-1}{\gamma+1}\,q_m^2 u(1-u)\right],$$

where

$$u \equiv \frac{1\ 67(CN)^{1\ 075}\,E}{W_m} \qquad (4\ 10)$$

This expression is valid only provided $u \leqslant 1$ The constant a is equal to 20 for lead and 23 for air or water We should regard (4 10) merely as a rather careful determination of the order of magnitude of $P(N)$ The expression (4 6) is of course much more accurate

We see at once from (4 10) that for a given E, $P(N)$ varies inversely as Z, i e the chance of a shower containing N electrons and positrons *with energy greater than* E being produced by a heavy particle is *inversely* proportional to the atomic number The reason for this is obvious, for the production of fast secondary electrons by collision is just proportional to the number of electrons, i e to Z, whereas the distances within which this production has to take place are determined by λ_0, the unit of length in the cascade theory for the material concerned which is inversely proportional to Z^2 The above result is, however, not of direct experimental interest

In the usual cosmic ray experiments we are interested in the relative frequency of showers containing different numbers of particles irrespective of their energy The lowest value we can give to E in (4 10) is the critical energy E_c. Substituting this in (4·10) we would get the probability of a shower containing N particles above the critical energy of the substance. To a rough approximation E_c may be put equal to $1600mc^2/Z$, so that in this case the factor outside the square brackets in (4·10) ceases to depend explicitly on Z. In other words, the frequency of *small showers for which $u \ll 1$ is independent of the atomic number of the material to a first approximation.* * *The average energy of the particles in a shower is however higher in lighter elements since this is roughly proportional to E_c.* More accurately, $E_c = 10^7$ e-volts in lead and $1·5 \times 10^8$ e-volts in air or water, and the value of a is also somewhat larger in lighter elements, so that the frequency of showers in lighter elements is also somewhat less than in heavier ones. *Large showers* will, however, be markedly less frequent in light elements due to the operation of the expression in square brackets in (4·10).

The size of the largest shower which occurs with any probability is determined by the fact that when $u = 1$ (34) vanishes. This gives

$$N_{\max} = 0·062\left(\frac{W_m}{E_c}\right)^{0·93} \qquad (4·11)$$

E_c being roughly inversely proportional to Z, it follows that $N_{\max}$ is crudely proportional to it. The frequency of small showers is very roughly inversely proportional to N.

All we have said above refers to the number of particles with energies above the critical energy, the total number of particles in a shower is roughly double this, as we shall deduce from the following general results of the cascade theory (Bhabha and Heitler 1937, fig. 4 and § 7, A). The number of particles in a shower with energy $> E$ is proportioned to $1/E$ provided $E \geqslant E_c$. The number in any energy range dE therefore varies as

---

* *Note added in proof.*—The recent experiments of Morgan and Nielsen (1937) show that the intensity of secondaries and showers in equilibrium after large thicknesses of absorber with penetrating particles is roughly the same in lead and iron, in agreement with our theory. This is not the case on the theory of Landau and Rumer, where only the process (2) is considered. (More exactly, Morgan and Nielsen find that the intensity in iron is 20 % higher. We believe that this discrepancy can be attributed to geometrical differences in the two cases, since the dimensions of the counter system are the same in both, whereas the characteristic units of length λ_0 are quite different in iron and lead.) They also observe typical transition effects on adding lead or iron after large thicknesses (274 g./cm.^{2}) of iron or lead respectively, the addition of lead causing a rapid increase followed by a decrease to the air lead curve, and the addition of iron to lead causing a rapid decrease followed by an increase to the air iron curve. These results are completely in accordance with our theory and result from the fact that although *the secondary intensities are roughly the same in all elements, the mean energy of the secondaries is much higher in lighter elements*, as we have already pointed out in the text.

dE/E^2 Multiplication does not continue below the critical energy E_c, so the number above E does not continue to increase as $1/E$ But the number of particles in a given energy range dE is roughly the same as at the critical energy, namely dE/E_c^2 Hence the total number of particles below E_c is just proportional to $1/E_c$ whereas the number above E_c is also proportional to E_c We thus get the following rough result of the cascade theory The number of particles in a shower with energies below the critical energy is roughly equal to the number above the critical energy, which is what we wished to deduce

In Table VI we give the values of $P(N)$ for lead and water assuming the constant C to be unity * The figures give the probability for a shower containing N particles above the critical energy The total number of particles in the shower is roughly twice this as we have already stated The figures show, what we have already said, that while the probability of shower production is less in lighter elements, it is nothing like as small as would follow from a Z^2 law This result seems to be in accordance with what little is known about shower production by the penetrating component We also see that the chance of a shower of some ten electrons and positrons accompanying the heavy particle is not very small, being of the order of 0 5% This seems to be of the right order of magnitude to agree with experiment Further, except for large showers, the figures are not very sensitive to the mass of the heavy particle

From these figures we may calculate the average number of particles accompanying the heavy particle This is what we have calculated fairly accurately in § 3 We thus get figures which are on the average about twice those of Table IV This is due to the fact that in calculating Table VI, we have put the constant C equal to unity whereas in fact we know that it is larger The comparison shows that C is nearly 2 and that hence the figures in Table VI are roughly too large by a factor 2 The agreement is as good as could be expected and confirms our statement that the formula (4 10) is a fairly accurate determination of the order of magnitude of the shower probabilities

Lastly we would remark that Table VI gives the shower probabilities due only to the first process mentioned at the beginning of this section The

---

* *Note added in proof* —We might have proceeded more accurately as follows Putting $E = E_c$ in (4 10), the mean number of particles of energy greater than E_c accompanying the heavy particle is just $\sum_N NP(N)$, which is a function of C This is just $N_{e>}$ of (3 2) By equating the two we can determine the value of C This is to be inserted in (4 10) in evaluating $P(N)$ We find that C is in the neighbourhood of 2 Due to the occurrence of C in u, this method would have given slightly smaller relative probabilities for larger showers than those of Table VI, in addition to reducing all the probabilities by roughly a half, as is stated in the text below

probabilities for the second process, i e for showers caused by quanta emitted by the heavy particles can be calculated similarly We shall not do this explicitly Indeed the expressions (4 2) and (4 3) show that for particles of mass $M \gg 10m$ the chance of emitting a hard quantum is negligible compared to the chance of producing an electron of the same energy Thus for particles of $M = 100m$ and protons the second process is negligible For particles of mass 5×10^6 e-volts the two probabilities are roughly equal, so that if we took the second process into account as well, all figures in Table VI for particles of 5×10^6 e-volts would be roughly doubled Moreover the second process would make larger showers relatively somewhat more probable since the emission of light quanta does not favour the lower energies as much as the production of fast secondaries

TABLE VI

$E_0 - Mc^2$	Mc^2	N	1	2	4	5	10	50
10^8 e volts	5×10^6 e volts	Lead	2 4	0 4	0 014	—	—	—
		Water	—	—	—	—	—	—
10^{10} e volts	5×10^6 e volts	Lead	4 7	1 5	0 58	0 40	0 12	$0\ 37 \times 10^{-2}$
		Water	2 5	0 72	0 25	0 16	0 034	—
	5×10^7 e volts	Lead	4 6	1 4	0 57	0 38	0 12	$0\ 34 \times 10^{-2}$
		Water	2 5	0 69	0 22	0 11	0 026	—
	Protons	Lead	2 9	0 6	0 08	0 025	—	—
		Water	—	—	—	—	—	—
10^{12} e volts	5×10^6 e volts	Lead	4 7	1 5	0 60	0 42	0 13	$0\ 54 \times 10^{-2}$
		Water	2 8	0 88	0 36	0 24	0 077	$0\ 29 \times 10^{-2}$
	5×10^7 e volts	Lead	4 7	1 52	0 60	0 42	0 13	$0\ 54 \times 10^{-2}$
		Water	2 8	0 88	0 36	0 24	0 077	$0\ 29 \times 10^{-2}$
	Protons	Lead	4 7	1 5	0 60	0 42	0 13	$0\ 53 \times 10^{-2}$
		Water	2 8	0 86	0 35	0 24	0 074	$0\ 26 \times 10^{-2}$

The figures give the probabilities per cent of the heavy particle being accompanied by a shower containing N particles above the critical energy The total number of particles in the shower is roughly twice this The upper figures in each row refer to lead, the lower figures to air or water *If showers started by emitted quanta be also taken into account, then the figures for particles of $M = 10m$, would be somewhat more than doubled, and the others would be unaffected*

5—THE CREATION OF HEAVY PARTICLES

In this section we will just discuss very briefly what effect the possibility of pair-creation of heavy particles would have on cosmic ray phenomena Since the pair-creation cross-sections vary roughly inversely as the square of the mass of the particle it follows that it is much more likely that an electron rather than a heavy particle should be created by a γ-ray Thus, if a heavy particle emit a quantum, the subsequent range of the quantum is almost entirely controlled by the probability of its creating electrons, and the number of quanta and electrons at any point is determined by the cascade process which follows The chance of a pair of heavy particles being created by one of the quanta emitted by a heavy particle is therefore a process of a higher order

The creation of heavy particles however has an effect on questions concerned with penetrating power Suppose a very high energy quantum or electron enters some material The number of particles and the ionization up to distances of 20 or 30 in the characteristic units λ_0 is determined by the usual cascade process In this process many quanta take part at some stage, and there is a finite though small chance of some heavy particles being created These, due to their low radiation loss would continue to penetrate to distances much larger than $30\lambda_0$ while the cascade electrons do not do so with any comparable probability The following question is therefore of interest Supposing a quantum of energy $h\nu$ enters a sheet of material, what is the number of heavy particles of mass M with energy greater than some arbitrary value E found at distances so large that the cascade process following on the original quantum has died out? The distances however have to be small compared to the range of the heavy particles Both these conditions can be fulfilled if M and E are large enough We have solved this problem very crudely, and give only the answer here The number of heavy particles with energy greater than E is of the order

$$\tau\left(1-\frac{E}{h\nu}\right)+\alpha\tau\frac{h\nu}{E}\left(1-\frac{E}{h\nu}\right)^2, \qquad (5\ 1)$$

where $$\tau = 0\ 6\left(\frac{m}{M}\right)^2$$

The order of magnitude of the expression (5 1) is determined by τ, and it is only valid when $\tau \ll 1$ for then the creation of heavy particles does not appreciably influence the cascade process

It is premature to draw any definite conclusions, but we wish to point out in passing that the above considerations may have some bearing on the following discrepancy between theory and experiment While the ordinary cascade theory predicts that an electron of 10^{10} e-volts at the top of the atmosphere will produce an entirely negligible number of electrons at sea level, Bowen, Millikan and Neher find a number of particles which is an order of magnitude larger, being about 1/500th of the number at the maximum A considerable fraction of these particles at sea level could be attributed to pair-creation in the atmosphere of heavy electrons with rest mass of the order of $10m$, if such pair-creation is possible at all This does not exclude the possibility that some heavy particles at sea level may have come in from outer space

In this connection we would remark that the well-known second maximum of the Rossi curve which occurs at 17 cm of lead shows that penetrating particles are also produced in the lead (cf Schmeiser and Bothe 1937)

6—GENERAL DISCUSSION AND CONCLUSIONS

In discussing the experimental evidence in the light of assumption (*b*) we have not limited ourselves to the hypothesis that only one new particle is concerned in cosmic radiation Indeed in the experimental evidence itself there are definite hints that one new particle alone may not suffice to explain all the facts We must therefore be prepared for the eventuality that a later and more complete theory may allow particles to exist whose rest masses may take on one of an infinite number of possible values of which only a few may turn out to be stable With this idea is connected the possibility of a particle changing its rest mass, the difference in the energy being radiated or communicated to some other particle in the immediate neighbourhood This change in the rest mass may be spontaneous, or caused by an external agency The former possibility is not very interesting as far as cosmic radiation is concerned, for if the probability is large, the change will take place before the particle reaches the earth, if small, then the chance of its taking place in the very short time taken by the particle in penetrating the earth's surface is also negligible We will therefore only consider the second possibility, and in particular the chance of a heavy electron changing its rest mass while moving very rapidly near a nucleus We consider the process in the system in which the heavy electron is at rest, so that the nuclear field acts like a superposition of quanta The change of rest mass with the emission of radiation then corresponds, as it were, to radiation induced by the presence of one of these quanta, the particle passing into a state of lower rest mass We of course do not ascribe any internal structure to the particle The emitted quantum would therefore be sent out in the direction in which the nucleus is moving, and hence in the system in which the nucleus is at rest would have very little energy indeed due to the action of the Lorentz transformation The particle would nevertheless have changed into a particle of smaller rest mass, and its subsequent behaviour would, therefore, be different The only chance of a large quantum being emitted in the system in which the nucleus is at rest is therefore if a direct collision takes place The cross-section for this is therefore at most the nuclear cross-section, i e of the order $(e^2/mc^2)^2$ and it may be very considerably smaller Since the process however depends on a direct collision, we might expect it to vary for different nuclei as Z rather than Z^2 This behaviour might be connected with the fact observed by Blackett and Wilson that for very high energy particles the energy losses in lead and aluminium are comparable Of course, the possibility of such radiation losses does not affect the ordinary Bremsstrahlung loss we have calculated above, but will be in addition to it

Finally, we wish to draw attention to a discrepancy of which not sufficient notice has been taken A comparison of the results of several investigators has shown, as has been put into direct evidence by Auger, Ehrenfest, Freon and Fournier (1937) that measurements of the absorption of cosmic radiation by change in the zenith angle of three counters in a plane are not in agreement with measurements in which the depth of the counters below the top of the atmosphere is changed *It directly follows that the absorption of the radiation is not a unique function of the thickness of atmosphere traversed* The discrepancy is not attributable to differences in intensity at the top of the atmosphere due to the earth's magnetic field since the results do not depend on the direction in which the zenith angle measurements are made Nor is this attributable to a transition effect as Blackett (1937*b*) suggests, since air and water have roughly the same atomic number, and further, geometry does not play a role when single track coincidences are measured These discrepancies however could be reconciled by the assumption that for some unknown reason the uppermost layer of the atmosphere absorbs more strongly than we should expect it to This would also be in agreement with the observation of Bowen, Millikan and Neher (1937) that the maximum of the absorption curve in the atmosphere occurs at about two-thirds of the distance from the top that the theory predicts

We may then sum up the results of the discussion of this paper as follows

1—The measurements of energy loss by Blackett and Wilson in lighter elements, if they are correct, are easily reconcilable with a "breakdown" theory in which the breakdown energy depends on the atomic number They can only be explained on a "new particle" theory by attributing a behaviour to the new particles which is not described entirely by present quantum mechanics Processes giving rise to large energy losses must be assumed which however vary as Z rather than Z^2, as for example a possible change in the rest mass of the particle (§ 1 C)

2—An analysis of the transition curves for large showers and bursts shows that a breakdown theory is unable to explain the shape of these curves, and allows one to conclude that at least some large bursts and showers must be entirely due to cascade processes in accordance with the theory (§ 1 B)

3—The latitude effect at sea level which extends to 50°, in conjunction with the direct proof by Bowen, Millikan and Neher that the cascade theory is correct in air up to 10^{10} e-volts demands the existence of new particles A breakdown theory cannot explain this sea level latitude effect since a breakdown below 10^{10} e-volts cannot be in question (§ 1 A)

4—A soft component consisting of electrons must accompany any energetic penetrating particle and *for heavy electrons of mass* $M \sim 100m$ is in air or water of the order of 7 % for particles of 10^{10} e-volts and of the order of 15 % for particles of 10^{12} e-volts The former figure is of the same order as the experimental findings of Auger and others In lead the soft component varies from about 15 to 30 % For $M \sim 10m$ the figures are roughly double these (§ 3)

5—The frequency of production of showers by a heavy particle does not depend on the atomic number for the smaller showers, though it is somewhat less in lighter elements The size of the largest shower which is produced with any reasonable probability by penetrating particles of a given energy is, however, proportional to Z The mean energy of the shower particles is higher in lighter elements for showers of the same size (§ 4)

6—The probability of a shower of N^{γ} particles being produced by a very energetic heavy particle is roughly inversely proportional to N^{γ} provided N is not too large, where γ is somewhat larger than 1 (§ 4)

7—The large number of particles observed at sea level by Bowen, Millikan and Neher may partly be heavy electrons of mass almost ten times the electron mass created in the atmosphere by the quanta emitted by electrons, though some of them may be heavy electrons which entered the atmosphere from outside (§ 5)

8—The discrepancy between the zenith angle and direct absorption measurements, and also the observation of Bowen, Millikan and Neher that the maximum of the atmospheric absorption curve occurs at two thirds of the theoretical thickness from the top, may both be explained by assuming that for some unknown reason the uppermost layer of the atmosphere absorbs more than we should expect it to, i e more like a substance of larger atomic number This may be due to the fact that the atoms in these layers are largely ionized (§ 6)

9—A breakdown theory based on limiting the acceleration of the electron in the rest system would carry with it not only a modification of the radiation formulae but also a diminution in the ionization loss According to Cosyns the ionization of cosmic ray particles seems to be less than the theoretical

SUMMARY

An analysis of the experimental data is carried through to show that a "breakdown" theory for radiation loss of electrons cannot explain (1) the latitude effect at sea level from latitudes of 35–50°, (2) the large number of particles found at sea level in the difference curves of Bowen, Millikan and Neher for charged particles of 10^{10} e-volts, (3) the shape of the transition curve for large bursts All these facts can be explained by assuming that the penetrating component consists of new particles with masses between those of the electron and proton But in order to explain the energy loss measurements of Blackett and Wilson, one must then assume that these particles suffer large energy losses in addition to the ordinary Bremsstrahlung which must vary in different substances as Z rather than Z^2, as for example a change in the rest mass of the particles

The radiation loss and the pair-creation cross-sections taking screening into account accurately for "heavy" electrons are calculated The frequency of the production of showers of different sizes by such heavy electrons as also the intensity of electrons in equilibrium with such particles forming a soft component are also calculated, and it is shown that though both these are somewhat larger in heavier elements, the variation is much less than a Z^2 law would give A comparison with experiment gives the mass of the penetrating particles as of the order $100m$

If such heavy electrons can be created as usual in pairs, then a part of the hard component at sea level could consist of heavy electrons of mass $10m$ created by the soft component in the upper atmosphere

It is shown that there are reasons to suppose that the uppermost layers of the atmosphere are more absorbing than one should expect from their mass This may be due to the fact that the atoms in these layers are largely ionized thus increasing the effective radiation and pair-creation cross-sections

REFERENCES

Alocco 1935 *Nature, Lond*, **135**, 96
Anderson and Neddermeyer 1936 *Phys Rev* **50**, 263
Arley and Bhabha 1937 In the Press
Auger and Ehrenfest 1934 *C R Acad Sci, Paris*, **199**, 1609
Auger, Ehrenfest, Freon and Fournier 1937 *C R Acad Sci, Paris*, **204**, 257
Auger, Ehrenfest, Freon and Grivet 1937 *C R Acad Sci, Paris*, **204**, 1797
Auger, Ehrenfest and Leprince Ringuet 1936 *J phys Radium*, **7**, 58

Auger and Leprince Ringuet 1934 *C R Acad Sci , Paris*, **199**, 785
Auger and Meyer 1933 *C R Acad Sci , Paris*, **204**, 572
Bethe 1934 *Proc Camb Phil Soc* **30**, 524
Bethe and Heitler 1934 *Proc Roy Soc* A, **146**, 83
Bhabha 1936 *Proc Roy Soc* A, **154**, 195
Bhabha and Heitler 1936 *Nature, Lond* , **138**, 401
— — 1937 *Proc Roy Soc* A, **159**, 432
Blackett 1937 *a* *Proc Roy Soc* A, **159**, 1
— 1937 *b* *Proc Roy Soc* A, **159**, 19
Blackett and Wilson 1937 *Proc Roy Soc* A, **160**, 304
Bøggild 1936 *Naturwissenschaften*, **24**, 280
Bowen, Millikan and Neher 1934 *Int Conf Nucl Physics*, London
— — — 1937 *Phys Rev* **52**, 80
Braddick and Gilbert 1936 *Proc Roy Soc* A, **156**, 570
Carlson and Oppenheimer 1937 *Phys Rev* **51**, 220
Carmichael 1936 *Proc Roy Soc* A, **154**, 223
Clay 1936 *Physica, Eindhoven*, **16**, 332
Clay, Gemert and Wiersma 1936 *Physica, Eindhoven*, **16**, 627
Compton and Stevenson 1934 *Phys Rev* **51**, 220
Ehrenberg 1936 *Proc Roy Soc* A, **155**, 532
Follett and Crawshaw 1936 *Proc Roy Soc* A, **155**, 546
Heisenberg 1936 *Z Phys* **101**, 533
Heitler 1937 *Proc Roy Soc* A, **161**, 261
Landau and Rumer 1937 *Nature, Lond* , **140**, 682
Messerschmidt 1936 *Z Phys* **103**, 27
Montgomery and Montgomery 1935 *Phys Rev* **47**, 318
Møller 1932 *Ann Phys , Lpz* , **14**, 531
Morgan and Nielson 1937 *Phys Rev* **52**, 564
Neddermeyer and Anderson 1937 *Phys Rev* **51**, 884
Rossi 1933 *a* *Nature, Lond* , **132**, 173
— 1933 *b* *Z Phys* **82**, 151
— 1934 *a* *Int Conf Nucl Physics*, London
— 1934 *b* *Ric Scient* **1**, VI, 559
Schmeiser and Bothe 1937 *Naturwissenschaften*, **25**, 669
Schwegler 1935 *Z Phys* **96**, 62
Street and Stevenson 1937 *Phys Rev* **51**, 1005
Street, Woodward and Stevenson 1935 *Phys Rev* **47**, 891
Woodward 1936 *Phys Rev* **49**, 711
Young 1937 *Phys Rev* **52**, 559

The Hydrolysis of the Methyl Halides

By E A Moelwyn-Hughes

(*Communicated by Eric K Rideal, F R S —Received* 19 *October* 1937)

An assumption common to the many current theories of chemical kinetics is that the velocity (k) of reaction is proportional to the Maxwell-Boltzmann factor $e^{-E/RT}$, which in simple cases represents the fractional number of molecules possessing an amount of internal energy not less than E The condition that reacting molecules must possess a critical energy in excess of the average value for all the molecules is, however, only one from a number of conditions which must in general be fulfilled before chemical change can take place The present development of the subject calls for quantitative information concerning these conditions, which have hitherto been discussed almost exclusively from the theoretical standpoint

When the complete expression for the velocity of reaction contains temperature-variable terms other than the Maxwell-Boltzmann factor, it is clear that the Arrhenius energy of activation (E_A), defined by the relation $E_A = RT^2\frac{\partial \ln k}{\partial T}$, cannot be constant To determine its variation with temperature requires, however, data of higher accuracy than those usually available, and our experimental knowledge on the problem is still meagre Perhaps the best example is the inversion of cane sugar, for which it has been found that $\frac{\partial E_A}{\partial T} = -98 \pm 4$ (Moelwyn-Hughes 1934) Values of the same order of magnitude must prevail also for disaccharides generally (Moelwyn-Hughes 1928) An important consequence of these results is that the true energy of activation is at least twice as great as the familiar but erroneous figures generally accepted, and that the corresponding Maxwell-Boltzmann factor in the expression for the velocity is incomparably less than one had imagined

The glycosides—particularly sucrose—are complicated structures, which, moreover, hydrolyse only in the presence of catalysts, and the questions arise as to how far high values of $-\partial E_A/\partial T$ are specific to changes suffered by large molecules, or to catalysis The experimental investigation of the uncatalysed hydrolysis of a suitable and simple molecule should allow of a decision Ethyl iodide presented itself as a favourable case

$$(C_2H_5I + H_2O \rightarrow C_2H_5OH + HI),$$

but its study is complicated by a side reaction ($C_2H_5I \rightarrow C_2H_4 + HI$) which is responsible for severe if only occasional explosions, and the work, for this reason and for others, was abandoned (Moelwyn-Hughes 1933)

We here describe experiments on the uncatalysed hydrolysis of three methyl halides, which we have found to be free from complications, and to proceed unimolecularly to completion at all accessible concentrations and temperatures The results in condensed form are shown in Table I, with energies in calories per gram-molecule, and velocity constants in reciprocal seconds These figures show beyond doubt that high values of $-\partial E_A/\partial T$ are not specific to the catalysed hydrolysis of glycosides, but hold also for the uncatalysed hydrolysis of methyl halides In both cases the values of $-\partial E_A/\partial T$ are of a higher order of magnitude than those found for reactions occurring between ions and polar molecules (Moelwyn-Hughes 1936*a*)

TABLE I

	CH_3Cl	CH_3Br	CH_3I
$k_{298\ 1}$	$1\ 99 \times 10^{-5}$	$3\ 57 \times 10^{-7}$	$6\ 87 \times 10^{-8}$
$k_{373\ 1}$	$1\ 08 \times 10^{-4}$	$1\ 31 \times 10^{-3}$	$4\ 47 \times 10^{-4}$
$E_{298\ 1}$	27,703	26,525	28,147
$E_{373\cdot 1}$	22,706	21,424	23,111
$\partial E_A/\partial T$	$-66\ 62$	$-68\ 01$	$-67\ 14$

It is too early yet to generalize, but we may say that a marked fall in the apparent energy of activation with rise in temperature characterizes both simple and complicated hydrolyses for which most accurate data are available

EXPERIMENTAL METHOD

Fig 1 illustrates the apparatus used for preparing the solutions Ordinary distilled water, which separate experiments showed to be as suitable as the purest conductivity water, is sucked into the 500 c c bulb *A*, where it is vigorously boiled at 20° C The vapour of the methyl halide, stored in the 3-litre bulb *B*, is led over anhydrous calcium chloride into the water in *A*, where it is allowed to bubble slowly for a day or two, pending saturation After measuring the partial pressure of the methyl halide, the solution is delivered into graduated reaction vessels by letting in air gently and at equal rates to both limbs of the bubbler, agitation of the solution leads to loss of solute vapour—a difficulty responsible for some of the irregularities found in previous work (Moelwyn-Hughes 1933) The sealed tubes are placed separately in pairs of concentric metal cylinders of the type described by Moelwyn-Hughes and Hinshelwood (1931) After immersion for a measured time in electrically regulated or in boiling-vapour thermostats, the contents of each tube are analysed for halide and hydrogen ions and for free halogen The end-point of each run corresponds to complete hydrolysis according to the equation

$$CH_3X + H_2O \rightarrow CH_3OH + H^+ + X^-$$

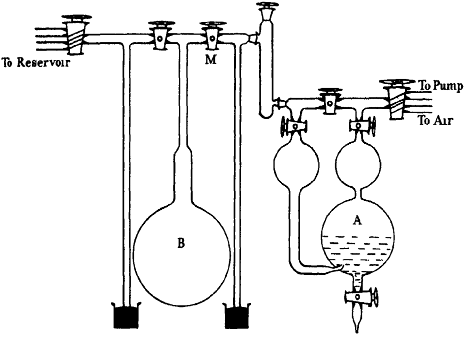


Fig 1

Due, however, to an inevitable escape of some solute vapour, the first tube of a series to be filled usually required rather more analytical reagent than the last, but the loss seldom exceeded 1%, and, as shown in Table II, was less in favourable circumstances

Table II—Standard reagent (c c) for representative end-point titrations

	CH_3Cl	CH_3Br	CH_3I
First tube filled	39 38	10 45	42 43
Last tube filled	38 90	10 38	42 12
First	67 40	24 96	43 84
Last	67 14	24 84	33 74
First	93 37	108 55	51 61
Last	92 43	108 50	51 58

That the end-points correspond to complete hydrolysis was shown by treating the hydrolysed solutions with alcoholic potash or with aqueous silver nitrate, when no further change could be detected. Previous work (1933) traced the formation of some free halogen—not exceeding 1%—quantitatively to atmospheric oxidation, hence in the present work no precautions were taken to exclude oxygen from the vapour above the solutions in the sealed tubes. In the case of methyl iodide, a tared solution made up to 0.3118*N* gave on analysis 0.3071*N* iodide, or 98.5%, and 0.88% molecular iodine. In the case of the other halides, the amount of free halogen formed was even less, and was ignored. The methyl halides, supplied by Kahlbaum, were found to be free from acid and uncombined halogen.

STATIC MEASUREMENTS

The distribution of methyl chloride and methyl bromide between the vapour phase and an aqueous solution was determined at 20° C., the particulars of a given experiment being as follows. Bulb *A* containing water saturated with methyl chloride had a total pressure of 278.5 mm. Hg. Subtracting the vapour pressure of pure water (17.5 mm.) and the hydrostatic pressure (1.5 mm.) due to a head of 2 cm. of water, the partial pressure, p, of methyl chloride in the gas phase is $278.5 - 19.0 = 259.5$ mm., assuming Dalton's law. The first tube of 20 c.c. required, after hydrolysis, 44.75 c.c. of 0.0114*N* sodium hydroxide to neutralize the acid generated, and the last tube required 44.05 c.c. Hence the concentration of solute is 25.31 mmol. per litre, and the distribution ratio (p/c) is 10.25 in these units. Other values found are reproduced in Tables III and IV. Occasionally, at the lower

TABLE III—DISTRIBUTION OF METHYL CHLORIDE AT 293.1° BETWEEN THE VAPOUR PHASE AND AN AQUEOUS SOLUTION

Partial pressure in the vapour phase (mm. Hg)	Concentration in solution (mmol./litre)	p/c
25.5	2.36	10.8
108.3	10.74	10.1
183	17.68	10.35
259.5	25.31	10.25
354	33.4	10.6
432.5	41.25	10.5
480	46.1	10.4
		Average = 10.4

pressures, high values of this coefficient were found, but repetition showed the tabulated figures to be correct, and the discrepancy to be due to loss of vapour, which has a relatively greater effect under these circumstances Within the region of pressures examined, Henry's law is obeyed Converting the concentrations in both phases to molecular units, we have $c_{\text{Solution}}/c_{\text{Vapour}} = 0\ 600$ in the case of methyl chloride, and 1 315 in the case of methyl bromide The corresponding energy terms,

$$w = RT \ln (c_{\text{Solution}}/c_{\text{Gas}}),$$

are -297 and $+159$ respectively These figures, while being useful for the kinetic purpose for which they were obtained, have no thermodynamic significance as they do not refer to equilibria

TABLE IV—DISTRIBUTION OF METHYL BROMIDE AT 293 1° BETWEEN THE VAPOUR PHASE AND AN AQUEOUS SOLUTION

p	c	p/c
42	8 85	4 75
210	43 9	4 78
335	70 7	4 74
350	73 2	4 78
472	99 9	4 73
497	104 2	4 77
498	104 5	4 76
		Average = 4 76

KINETIC MEASUREMENTS

Unimolecular velocity coefficients were determined by the usual methods from the equation

$$k = \frac{1}{t_2 - t_1} \ln \frac{T\infty - T_1}{T\infty - T_2},$$

where T stands for the titre at time t sec Determination of halide by Volhard's method, using standard silver nitrate and potassium thiocyanate with ferric alum as indicator, gave results in agreement with estimates of acidity using litmus For example, Exp 12 of the methyl bromide series gave $k = 1\ 84_0 \times 10^{-4}$ sec $^{-1}$ by the first method, while Exp 2 performed some weeks previously at the same temperature gave $1\ 83_4 \times 10^{-4}$ by the second method The two procedures were adopted indiscriminately, except in the case of methyl chloride, where the estimation of hydrogen ion is less laborious than the other Some specimen results are shown in Tables V, VI and VII The positive titre for $t = 0$ in Table V indicates some delay

TABLE V—METHYL CHLORIDE IN WATER AT 357 30°

t min	T_t, c c 0 00950N NaOH per 20 c c sample	$k \times 10^5$ sec $^{-1}$
0	0 22	—
50	5 00	2 74
134	11 04	2 70
180	13 74	2 76
288	20 81	2 70
420	27 86	2 75
560	33 20	2 70
830	40 77	2 65
∞	55 53	—

Average 2 71

Duplicate 2 67

$k = 2\ 69 \times 10^{-5}$

TABLE VI—METHYL BROMIDE IN WATER AT 330 70°

t min	T_t, c c 0 02020N $AgNO_3$ per 10 c c sample	$k \times 10^4$ sec $^{-1}$
0	0	—
35	2 43	2 43
88	5 92	2 46
109	7 13	2 43
140	9 11	2 48
256	15 42	2 49
300	17 30	2 45
412	22 08	2 46
482	24 75	2 47
∞	49 49	—

Average 2 46

Duplicate 2 48

$k = 2\ 47 \times 10^{-4}$

TABLE VII—METHYL IODIDE IN WATER AT 321 23°

t min	T_t, c c 0 02018N $AgNO_3$ per 20 c c sample	$k \times 10^6$ sec $^{-1}$
0	0	—
524	2 48	1 95
865	4 08	1 94
1475	6 83	1 95
1953	9 00	1 99
2770	12 03	1 96
2932	12 85	2 00
2735	15 22	1 93
∞	43 62	—

Average 1 96

Duplicate 1 90

$k = 1\ 93 \times 10^{-6}$

between charging the tubes and starting the run Results to be described in the following paragraph suggest the possibility of a slight variation in k during the course of a single experiment No such change could, however, be detected, where drifts occurred, they were as often upwards as downwards, and are probably due to experimental error

THE DILUTION EFFECT

The extent to which the rate of hydrolysis is influenced by the dilution is illustrated in Table VIII, where the initial concentration of methyl bromide has been varied ninefold, which is the limit attainable by the present apparatus and technique

TABLE VIII—THE VARIATION OF k WITH DILUTION

T ° abs	Initial concentration of methyl bromide g mol /litre	$k \times 10^4$ sec $^{-1}$
363 40 ± 0 06	0 0122	5 28 ± 0 10
363 33 ± 0 06	0 0303	5 58 ± 0 08
363 12 ± 0 04	0 1045	6 06 ± 0 10

The reaction was in each case followed until at least 75% of the total change The results are approximately expressible by a square-root law ($k \times 10^4 = 5\ 00 + 3\ 3c^{\frac{1}{2}}$), but supplementary evidence points to a limiting velocity constant for the saturated solution The results leave no doubt as to the kinetic order of reaction

THE INFLUENCE OF TEMPERATURE

With a fixed initial concentration of solute, the velocity of hydrolysis has been determined over a temperature range sufficient, in each case, to alter the rate by a factor of at least one thousand The results do not conform with an equation of the Arrhenius type, but may be represented, as shown in Tables IX, X and XI, by the following formulae

$$\text{Methyl chloride} \quad \log_{10}k = 110\ 223 - \frac{10{,}403}{T} - 33\ 559 \log_{10}T,$$

$$\text{Methyl bromide} \quad \log_{10}k = 112\ 656 - \frac{10{,}236}{T} - 34\ 259 \log_{10}T,$$

$$\text{Methyl iodide} \quad \log_{10}k = 111\ 859 - \frac{10{,}534}{T} - 33\ 821 \log_{10}T$$

These expressions are valid as first approximations, and have been used to obtain the values of E_A and $\partial E_A/\partial T$ given in Table I The most accurate set of data refer to methyl bromide, on account of its relatively great solubility Detailed analysis of all the results suggest that $-\partial E_A/\partial T$ passes through a maximum within the experimental temperature region, but from the form of the above empirical equations, the second differential coefficient of E_A with respect to temperature cannot be estimated with accuracy The last numerical constant in these expressions (and consequently $\partial E_A/\partial T$) is to be regarded as the average value for the explored temperature interval In the case of methyl bromide, it varies from the actual values in the different temperature regions by about ± 10% of the mean value The variations in the other cases are higher

TABLE IX—METHYL CHLORIDE (0 0376*N*)

T	k_{observed}	k_{formula}
314 78	$2\ 30 \times 10^{-7}$	$2\ 21 \times 10^{-7}$
326 24	$9\ 58 \times 10^{-7}$	$9\ 60 \times 10^{-7}$
333 11	$2\ 16 \times 10^{-6}$	$2\ 18 \times 10^{-6}$
341 28	$5\ 43 \times 10^{-6}$	$5\ 40 \times 10^{-6}$
348 99	$1\ 22 \times 10^{-5}$	$1\ 20 \times 10^{-5}$
357 30	$2\ 69 \times 10^{-5}$	$2\ 70 \times 10^{-5}$
373 37	$1\ 08 \times 10^{-4}$	$1\ 10 \times 10^{-4}$
383 98	$2\ 48 \times 10^{-4}$	$2\ 54 \times 10^{-4}$

TABLE X—METHYL BROMIDE (0 0949*N*)

T	k_{observed}	k_{formula}
290 13	$1\ 07 \times 10^{-7}$	$1\ 04 \times 10^{-7}$
308 83	$1\ 65 \times 10^{-6}$	$1\ 65 \times 10^{-6}$
319 50	$6\ 71 \times 10^{-6}$	$6\ 57 \times 10^{-6}$
330 70	$2\ 47 \times 10^{-5}$	$2\ 46 \times 10^{-5}$
335 02	$3\ 91 \times 10^{-5}$	$3\ 94 \times 10^{-5}$
350 27	$1\ 84 \times 10^{-4}$	$1\ 84 \times 10^{-4}$
363 38	$6\ 06 \times 10^{-4}$	$6\ 12 \times 10^{-4}$
373 18	$1\ 28 \times 10^{-3}$	$1\ 31 \times 10^{-3}$

TABLE XI—METHYL IODIDE (0 0467*N*)

T	k_{observed}	k_{formula}
308 87	$3\ 43 \times 10^{-7}$	$3\ 53 \times 10^{-7}$
321 23	$1\ 93 \times 10^{-6}$	$1\ 92 \times 10^{\ 6}$
323 69	$2\ 70 \times 10^{-6}$	$2\ 64 \times 10^{-6}$
328 72	$4\ 81 \times 10^{-6}$	$4\ 925 \times 10^{-6}$
333 88	$8\ 72 \times 10^{-6}$	$9\ 08 \times 10^{-6}$
342 92	$2\ 49 \times 10^{-5}$	$2\ 50 \times 10^{-5}$
363 06	$1\ 80 \times 10^{-4}$	$1\ 84 \times 10^{\ 4}$
372 92	$4\ 31 \times 10^{-4}$	$4\ 34 \times 10^{-4}$

Discussion

The principal feature of these results is that the temperature coefficient of the apparent energy of activation is as great as -67 for the most simple cases of uncatalysed hydrolyses hitherto investigated No single theory yet advanced can consistently account for this magnitude, and it appears that many of the hypotheses independently furthered in various directions must be co-ordinated to furnish a satisfactory solution of the problem Before briefly referring to these, we may make some inferences as to the mechanism of hydrolysis on the basis of the present experimental facts

The methyl halide molecule is capable of nine internal motions, most of which, from the known molecular frequencies, are fairly stiff vibrations, incapable at ordinary temperatures of contributing appreciably to the specific heat term To explain the present observations we must consequently look to certain extramolecular effects, such as the rotation of the solute molecule as a whole in its solvent envelope, or to the weak vibrations which characterize the solute-solvent contacts The measurable step in hydrolysis requires the participation of many bonds, exceeding in number those which are found within a single solute molecule The reaction is thus a multimolecular one, although there is no objection to regarding it as the unimolecular decomposition of a cluster of solvent molecules with a solute molecule at the centre The conclusion tallies well with the fact that the products of reaction are solvated ions with a pseudocrystalline structure The halide ion retains this structure, but the methyl ion replaces a hydrogen ion from a water molecule

The decomposition of the alkyl halides in the gaseous phase (Vernon and Daniels 1932) requires a greater apparent energy of activation than the true energy of activation (*ca* 48,000) for the hydrolysis The difference is great enough to render improbable, though not to exclude, the mechanism of a rate-determining formation of free radicals in solution, followed by rapid polar changes, including ionic solvation

Another mechanism, provisionally ignored pending further work, is the simultaneous approach of a hydroxyl ion and a hydrogen ion from opposite sides of the methyl halide molecule (cf Ogg and Polanyi 1935, Moelwyn-Hughes 1936*b*)

Accepting the rate-determining step to be the simultaneous attack of the solute molecule by many of the solvent molecules which surround it, we may note three effects which can contribute to the fall in the Arrhenius energy of activation (1) Let some of the energy reside in a number, F, of bonds which are assumed to oscillate classically and to be feebly coupled, then $E_A = \text{constant} - F\ RT$ (Hinshelwood 1926, Fowler and Rideal 1927) It is improbable that the whole of the observed effect is due to this distribution of energy, if it were, the activated cluster would have an average life of 1 sec (when $F = 34$, $T = 373\ 1$, $E = 47{,}600$, activated fraction $= 1\ 86 \times 10^{-4}$) Moreover, in cases such as the inversion of cane-sugar ($F = 49$, $E = 58{,}200$, $T = 298\ 1$, activated fraction $= 1\ 74 \times 10^{-7}$), the supply of activated complexes is barely sufficient to maintain a stationary concentration It seems evident, however, that the distribution of the activation energy among a number of degrees of freedom accounts for some of the observed effect, and yields a theoretical expression which is formally compatible with experiment This opens up the question of how far the actual motions resulting from the interaction of solute and solvent molecules may be treated as classical oscillations (2) Let the solute and solvent molecules have electrical moments of magnitude μ_A and μ_s respectively, and an average separation $\bar{r}_0$ Assuming the critical separation, which differs from $\bar{r}_0$, to be independent of temperature, and considering one solute molecule to make contact with x solvent molecules, we have

$$E_A = \text{constant} + \frac{x\mu_A\mu_s}{\bar{r}_0^3}\left\{1 + 3\left(\frac{\partial \ln \bar{r}_0}{\partial \ln T}\right)_P\right\} \quad \text{(Moelwyn Hughes 1937)}$$

Hence the existence of dipole-dipole attractions also partly accounts for the observed fall in the Arrhenius constant, though the motion implied cannot be a simple harmonic one (3) The acquisition and appropriate distribution of internal energy by the reacting cluster and the adjustment of potential energy are not the only conditions which must be fulfilled before chemical change can take place The internally activated solute molecule must turn round in the sheath of solvent molecules to find a position which, from the purely steric as well as from the potential-energy consideration, is most favourable to reaction To do so entails the expenditure of energy, which varies with the temperature, and may be estimated from the average speed of rotation of a solute molecule Using Einstein's expression for a spherical molecule of radius r in a medium of viscosity η, we have $\bar{\nu} = kT/4\pi r^3\eta$ This effect accounts for a partial $\partial E_A/\partial T$ of about -15

Not one of the three expressions discussed here can be regarded as quantitatively applicable to the systems under investigation, but they all indicate directions along which further progress may be sought

The average decrease in molecular heat attending the dissociation of fifteen electrolytes, chiefly acids and bases, in aqueous solution at 298 1° is $-38\ 8$ The reverse process, i e the combination of ions to form undissociated molecules, is known to be accompanied by much smaller specific heat changes, hence the ionization of electrolytes in water may be included, along with the hydrolysis of glycosides and alkyl halides, in that category of chemical changes for which the equation of Arrhenius is not applicable These examples are experimentally more reliable than the dubious instances cited by La Mer (1933), with whose theoretical treatment, however, we are in agreement While this work was in progress, two further instances have come to light of reactions which probably belong to this category They are the mutarotation of glucose, for which $\partial E_A/\partial T = -27\ 0 \pm 1\ 6$ (Smith and Smith 1937), and the decomposition of oxalic acid in water, for which a much higher value holds (Dinglinger and Schroer 1937) It is rather surprising that no such effect has yet been detected in the numerous hydrolyses and alcoholyses of the substituted methyl halides (e g Dawson and Pycock 1936, Hughes and Ingold 1937) Though $\partial E_A/\partial T$ in solvents other than water may well be relatively small, the "fall effect" must be regarded in these instances as unrevealed rather than absent

As far as we are able to generalize, we may say that (1) The inapplicability of the Arrhenius equation is most marked in reactions involving solute and solvent These reactions have for some time been regarded as somewhat special cases, requiring a modification of the Arrhenius law (Moelwyn-Hughes 1933), hence the present discoveries, while demanding an adjustment of ideas, need not distort the general presentation of the kinetics of reactions in solution (2) The temperature coefficient of the apparent energy of activation is always negative in the regions explored

The present work was carried out in collaboration with Professor E K Rideal, F R S, for whose advice and guidance I am once more greatly indebted

SUMMARY

For a chemical reaction proceeding with measurable rate, the most significant single constant is the energy of activation, found by the familiar method of Arrhenius The value obtained is, in the majority of known cases, a true constant over the accessible temperature regions In the case of the catalysed hydrolysis of some complicated molecules, namely the glycosides, it has, however, been found (1934) that the energy of activation falls markedly with a rise in temperature As the result carried significant theoretical consequences, it became necessary to give an experimental answer to the question as to whether this fall was specific to the acidic hydrolyses of glycosides, or of a more general character In the present

work, we have investigated the uncatalysed and uncomplicated hydrolysis of some very simple molecules, namely the methyl halides, and have found that in these cases also the Arrhenius constant falls as the temperature is raised

The fall effect seems to be most pronounced for, if not confined to, reactions involving collisions between solute and solvent molecules, and as far as can be judged is always negative The conclusion is confirmed by independent collateral evidence

At least three causes are recognized as responsible for the total effect They are (1) the distribution of energy among a number of oscillators, (2) the electrostatic interaction of the polar molecules of solute and solvent and (3) the frequency of the rotation of a solute molecule in the solvent atmosphere

REFERENCES

Dawson, H M and Pycock, E R 1936 *J Chem Soc* pp 153–8

Dinglinger, A and Schroer, E, 1937 *Z phys Chem* A, **179**, 401–26

Fowler, R H and Rideal, E K 1927 *Proc Roy Soc* A, **113**, 570

Hinshelwood, C N 1926 *Proc Roy Soc* A, **113**, 230

Hughes, E D, Ingold, C K and collaborators 1937 *J Chem Soc* pp 1177–1291

La Mer, V K 1933 *J Chem Phys* **1**, 289–96

Moelwyn Hughes, E A 1928 *Trans Faraday Soc* **24**, 309–21

— 1929 *Trans Faraday Soc* **25**, 503–20

— 1933 *J Chem Soc* pp 1576–80

— 1934 *Z phys Chem* B, **26**, 281–7

— 1936*a* *Proc Roy Soc* **157**, 667

— 1936*b* *Acta Phys U R S S* **4**, 173–224

— 1937 *Trans Faraday Soc*, Manchester Symposium (in the Press)

Moelwyn Hughes, E A and Hinshelwood, C N 1931 *Proc Roy Soc* A, **131**, 177–98

Ogg, R A and Polanyi, M 1935 *Trans Faraday Soc* **31**, 604

Smith, G F and Smith, M C 1937 *J Chem Soc* pp 1413–20

Vernon and Daniels, F 1932 *J Amer Chem Soc* **54**, 1870

Significance tests for continuous departures from suggested distributions of chance

By Harold Jeffreys, F R S

(*Received* 1 *November* 1937—*Revised* 18 *November* 1937)

1—The problem to be considered is as follows A trial law gives the chance of an observation in a given range of an argument as a continuous function of the argument, for instance, the suggested distribution of chance may be uniform or normal We wish to know whether a set of observations indicates any departure from this law, thus for a suggested uniform distribution we may want to test a linear or a harmonic departure, and for a suggested normal one we may want to test a symmetrical departure giving excesses or deficiencies at the tails

In the first place we consider problems where the trial hypothesis q is that the distribution of chance is uniform, we can then choose a linear function t of the argument x, so that t will be 0 at the lower and 1 at the upper limit The chance of an observation falling in a range dx is dt, and that of n observations in specified ranges is $\Pi(dt)$, provided that they are independent

On the extended hypothesis $\sim q$ we can use the same t as a parameter, but the chance will be $p(t)\,dt = \{1+af(t)\}\,dt$, where $f(t)$ is a specified function and a a parameter to be found Since the total chance must be 1 we must have

$$\int_0^1 f(t)\,dt = 0, \tag{1}$$

and without loss of generality we can take

$$\int_0^1 f^2(t)\,dt = 1 \tag{2}$$

Then the chance of all the observations, given $\sim q$, is $\Pi[\{1+af(t)\}\,dt]$

If two functions $f(t)$ and $g(t)$ naturally arise at once, as for a harmonic variation, the corresponding form will be $\Pi[\{1+af(t)+bg(t)\}\,dt]$, with

$$\int_0^1 f(t)\,dt = \int_0^1 g(t)\,dt = 0, \tag{3}$$

$$\int_0^1 f^2(t)\,dt = \int_0^1 g^2(t)\,dt = 1, \quad \int_0^1 f(t)g(t)\,dt = 0 \tag{4}$$

2—The problem of the introduction of one new function $f(t)$ suggests comparison with the tests already given for the introduction of new functions to express a series of measures (Jeffreys 1936), and for the consistency of a sampling ratio with a predicted value (Jeffreys 1937) It differs from the former, because the standard error of any determination from the data, given the hypothesis of independence, cannot be less than a minimum fixed by the number of observations, in the former problem the measures were supposed to have an unknown standard error, which might be indefinitely small, and our problem was to say how much of the outstanding variation was to be attributed to the random error and how much, if any, to systematic variation * The latter test could be adapted to the present problem, since the trial hypothesis predicts the chances of an observation in the ranges where $f(t)$ is negative or positive If then we combine all observations in these ranges we can compare their numbers with the predicted ratio by means of the test for sampling ratios This method, however, would involve a sacrifice of information if $f(t)$ is a continuous function An excess of observations where $f(t)$ is small would receive as much weight as one where it is large, but the latter would be more convincing evidence for the presence of $f(t)$ Neither test would therefore be satisfactory as it stands The former resembles the present problem, since it allows for $f(t)$ being a continuous function, the latter because it allows for the fact that the standard error cannot be indefinitely small, but each fails where the other succeeds A special discussion is therefore needed

The possible values of a are limited by the fact that $p(t)$ cannot be negative Thus if $-1/\alpha$ and $1/\beta$ are the extreme negative and positive values of $f(t)$, a must lie between $-\beta$ and $+\alpha$ Within this range we can take its prior probability uniformly distributed It may be desirable, however, to consider an alternative statement of $\sim q$ If q is far from the truth, other functions than $f(t)$ may occur in the chance The extreme departure possible in some problems may be that the chance is zero wherever $f(t)$ is negative and uniform wherever it is positive, or conversely It will no longer be expressible in the form $1+af(t)$, but it will be expressible with the addition of other terms orthogonal to $f(t)$, whose coefficients we may suppose to be small compared with a when a is itself small Then if the total length of the ranges where $f(t)$ is positive is λ, the maximum positive value of a will be given by

$$a = \int_\lambda p(t)f(t)\,dt \tag{1}$$

---

* The statement of the problem in this way is virtually given in Fisher's convenient term "The Analysis of Variance"

and the maximum negative one by

$$a = \int_{1-\lambda} p(t) f(t)\, dt, \tag{2}$$

$p(t)$ being assigned for the respective cases We denote the range of the permissible values of a by c

In the case where $f(t)$ is a linear function it must be $\sqrt{(12)}\,(t-\tfrac{1}{2})$ The extreme values of a in the first case will be $\pm 1/\sqrt{3}$ The maximum possible in the second case is given by

$$a = \int_{\frac{1}{2}}^{1} 2\sqrt{(12)}\,(t-\tfrac{1}{2})\, dt = \tfrac{1}{2}\sqrt{3} \tag{3}$$

Thus c will be $2/\sqrt{3}$ and $\sqrt{3}$ in the respective cases

Proceeding as usual, we have

$$P(q \mid h) = \tfrac{1}{2}, \quad P({\sim} q, da \mid h) = \frac{1}{2}\frac{da}{c}, \tag{4}$$

$$P(\theta \mid qh) = \Pi(dt), \tag{5}$$

$$P(\theta \mid {\sim} q, da, h) = \Pi[\{1 + af(t)\}\, dt], \tag{6}$$

$$P(q \mid \theta h) \propto 1, \tag{7}$$

$$P({\sim} q, da \mid \theta h) \propto \Pi\{1 + af(t)\}\frac{da}{c} \tag{8}$$

Then the posterior probability density for a, on ${\sim} q$, is e^{ϕ}/c, where

$$\phi = \Sigma \log\{1 + af(t)\}, \tag{9}$$

$$\phi' = \Sigma \frac{f(t)}{1 + af(t)}, \quad \phi'' = -\Sigma \frac{f^2(t)}{\{1 + af(t)\}^2} \tag{10}$$

The maximum is at $a = a_0$, where $\phi'(a_0) = 0$, if a_0 is small (as it will be whenever a test is needed) this is

$$a_0 = \frac{\Sigma f(t)}{\Sigma f^2(t)} = \frac{1}{n}\Sigma f(t), \tag{11}$$

nearly, where n is the number of observations Then

$$\phi(a_0) = \tfrac{1}{2}a_0^2\,\Sigma f^2(t) = \tfrac{1}{2}na_0^2, \tag{12}$$

$$\phi''(a_0) = -n, \tag{13}$$

nearly Hence

$$P(\sim q \mid \theta h) \propto \frac{1}{c}\int \exp(\tfrac{1}{2}na_0^2)\exp\{-\tfrac{1}{2}n(a-a_0)^2\}\,da = \sqrt{\left(\frac{2\pi}{n}\right)}\frac{1}{c}\exp(\tfrac{1}{2}na_0^2), \quad (14)$$

$$K = \frac{P(q \mid \theta h)}{P(\sim q \mid \theta h)} = \sqrt{\left(\frac{n}{2\pi}\right)}c\exp(-\tfrac{1}{2}na_0^2) \quad (15)$$

The term in $f(t)$ will therefore be supported if a_0 is such as to make this less than 1 The standard error of a_0, in this notation, is $n^{-\frac{1}{2}}$, so that the exponential factor has the usual form $\exp(-\tfrac{1}{2}\chi^2)$

The following table, for various values of n, gives K for $a_0 = 0$ for the two values of c, and the values of χ^2 and $a_0 n^{\frac{1}{2}}$ that make $K = 1$ For comparison we may notice that Fisher's (1936, Table III) 5 % and 1 % limits, for one degree of freedom, are at $\chi^2 = 3\ 84$ and 6 64, the former would agree in the first case at about 200 observations, the latter at about 4000 In the second case the agreements would come at about 100 and 1700 observations His test, of course, does not mean quite the same thing, it says when an observed result would be surprising on hypothesis q, whereas mine, for the larger numbers of observations, may admit this and yet say that it would be still

n	K		$a_0 n^{\frac{1}{2}}$		χ^2	
5	1 03	1 55	0 25	0 94	0 06	0 88
10	1 46	2 19	0 87	1 25	0 76	1 57
20	2 06	3 09	1 20	1 50	1 45	2 26
50	3 26	4 89	1 54	1 78	2 36	3 17
100	4 61	6 92	1 75	1 97	3 06	3 87
200	6 51	9 76	1 94	2 13	3 75	4 56
500	10 31	15 46	2 16	2 34	4 67	5 48
1,000	14 6	21 9	2 32	2 48	5 36	6 17
2,000	20 6	30 9	2 46	2 62	6 05	6 86
5,000	32 6	48 9	2 64	2 79	6 97	7 78
10,000	46 1	69 2	2 77	2 86	7 66	8 19

more surprising on $\sim q$ In any event cases where the observed a_0 would come in the disputable region would be expected to be rare if *either* of the hypotheses compared was correct, and some third alternative may suggest itself The low critical values for $n = 5$ and 10 might be altered by a more accurate solution, but their correct interpretation is that with so few observations the result will be indecisive in any case Thus if we take $a_0 = 1/\sqrt{3}$ as the extreme value consistent with the first case, we have for $n = 10$

$$a_0\sqrt{n} = 1\ 83, \quad \exp(-\tfrac{1}{2}\chi^2) = 0\ 19, \quad K = 0\ 28$$

Thus a set of 10 observations all as favourable as possible to $\sim q$ make K about 0 3, there would be about 3 5 to 1 odds on $\sim q$ on the data

The case $f(t) = -1$ for $t < \frac{1}{2}$, $f(t) = 1$ for $t > \frac{1}{2}$ would mean that on q there is an even chance that $t < \frac{1}{2}$, but that on $\sim q$ there is a difference between the chances of positive and negative t, these chances themselves being uniformly distributed In this case $c = 2$ and

$$K = \left(\frac{2n}{\pi}\right)^{\frac{1}{2}} \exp(-\tfrac{1}{2}\chi^2) \qquad (16)$$

This may be compared with the test of departure from a predicted ratio in a sample (Jeffreys 1937)

$$K = \left\{\frac{x+y}{2\pi p(1-p)}\right\}^{\frac{1}{2}} \exp(-\tfrac{1}{2}\chi^2), \qquad (17)$$

which reduces to (16) when $x+y = n$, $p = \frac{1}{2}$ It has already been noticed that in this case no very decisive result is obtained if the number of observations is less than about 7, even when they all agree For the linear form of $f(t)$ in the present problem the ratio of the standard error of a to its extreme possible value is larger, and the extreme of χ^2 correspondingly smaller, and decisive results need more observations

3—*Grouping of data* In the above work the estimate of the unknown a is practically identical with that given by the method of maximum likelihood, using the observations in detail Any preliminary grouping involves some loss of accuracy, in problems of pure estimation this loss may not be important, but when the significance of the parameters is not yet settled, the increase of the standard error may reduce the contribution to χ^2 from a given systematic variation by enough to bring it below the critical value when an accurate determination would give a significant χ^2 On the other hand if the standard error or, for harmonic variations, the equivalent Schuster criterion, is adopted from a formula adapted to an accurate discussion, and then compared with amplitudes found by grouping, large random amplitudes will appear to occur more often than they should, and may be accepted as genuine Grouped data, treated correctly, will therefore sometimes fail to detect variations that more accurate methods would find, treated incorrectly, will sometimes give systematic variations that more accurate methods would show to be spurious

Nevertheless grouping sometimes saves much labour, and it is desirable to have tests adapted to it If we take first the case of a linear variation, the coefficient may be found by comparing the numbers of observations in the ranges $t = 0$ to p and $t = 1-p$ to 1 If we call these n_{-1} and n_1, n_0 being the number in the intermediate range, and the true coefficient of $t-\frac{1}{2}$ is a, the expectations for the three ranges are

$$np\{1-\tfrac{1}{2}a(1-p)\}, \quad n(1-2p), \quad np\{1+\tfrac{1}{2}a(1-p)\}$$

Hence our estimate will be

$$np(1-p)\,a = n_1 - n_{-1} \tag{1}$$

We therefore need the standard error of $n_1 - n_{-1}$ Now the sampling error of n_0, if known, would tell us nothing about the sign of the error of $n_1 - n_{-1}$, the error of which is the difference of two equal and opposite errors, therefore $\sigma(n_1 - n_{-1})$ is twice the standard error of n_1 given that $n_1 + n_{-1} = 2np$ nearly, and the latter is $\frac{1}{2}\sqrt{(n_1 + n_{-1})}$ Thus

$$np(1-p)\,a = n_1 - n_{-1} \pm (2np)^{\frac{1}{2}} \tag{2}$$

Thus the uncertainty of a will be made a minimum by taking p so that $(1-p)\,p^{\frac{1}{2}}$ is a maximum, and therefore $p = \frac{1}{3}$ This result was obtained for measures by Eddington It makes

$$\sigma_a^2 = 27/2n \tag{3}$$

If we took $p = \frac{1}{4}$, we should have

$$\sigma_a^2 = 16/n \tag{4}$$

The present a is $1/\sqrt{12}$ times the a_0 of 2 (11), which had a standard error $1/\sqrt{n}$, hence the most accurate possible determination of a would have a standard error given by

$$\sigma_a^2 = 12/n \tag{5}$$

The comparison of the first and last thirds of the range therefore approaches the best value in accuracy, and nearly three times as closely as the comparison of the first and second halves of the range In Fisher's terminology, its efficiency is 89 %, that of the comparison of halves is 75 %

If this method is used we shall need an estimate of the outside factor in a significance test, the contribution to χ^2 from a given a will be 8/9 of that in the accurate solution The outside factor will be the same as for testing an even chance in a sample of number $2n/3$, and this is $\sqrt{(4n/3\pi)}$ The accurate solutions gave, on different hypotheses about the extreme types of variation admissible, the outside factors $\sqrt{(2n/3\pi)}$ and $\sqrt{(3n/2\pi)}$ The critical values of χ^2 will therefore be slightly less than for the latter solution

Similarly, if a Fourier term is $a\cos 2\pi t$, we may try to estimate it by comparing numbers of observations between $t = -p$ and p and between $\frac{1}{2}-p$ and $\frac{1}{2}+p$ The expectations for three ranges analogous to the above will be

$$2n\left(p+\frac{a}{2\pi}\sin 2\pi p\right),\quad n(1-4p),\quad 2n\left(p-\frac{a}{2\pi}\sin 2\pi p\right)$$

Hence our estimate of a will be given by

$$\frac{2na}{\pi}\sin 2\pi p = n_{-1}-n_1, \tag{6}$$

and the standard error of $n_{-1}-n_1$ is twice $\sigma(n_1)$ given $n_{-1}+n_1 = 4np$, namely

$$\sigma(n_{-1}-n_1) = 2\sqrt{(np)} \tag{7}$$

Hence
$$\sigma(a) = \pi\left(\frac{p}{n}\right)^{\frac{1}{2}}\operatorname{cosec} 2\pi p \tag{8}$$

This has a minimum for $2\pi p = 66^\circ\,47'$, namely

$$\sigma^2(a) = 2\ 168/n \tag{9}$$

The convenient value $2\pi p = 60^\circ$ gives

$$\sigma^2(a) = 2\ 193/n, \tag{10}$$

and therefore a negligible loss of accuracy $2\pi p = 90^\circ$ gives

$$\sigma^2(a) = 2\ 467/n, \tag{11}$$

so that proceeding right up to the vanishing of $\cos 2\pi t$ gives an appreciable loss The standard error of a, as found from a complete Fourier analysis, would be given by

$$\sigma^2(a) = 2\ 00/n \tag{12}$$

The effect of finding a by comparison of opposite ranges of 120° is therefore only to increase its standard error by about 5 % (Incidentally it eliminates all harmonics whose arguments are multiples of two or three times that of the fundamental) The respective efficiencies are 92 %, 91 %, 81 %, 100 % With the rule (10) the outside factor in a significance test would again be $\sqrt{(4n/3\pi)}$, if the occurrence of a cosine does not suggest the presence of the corresponding sine and require both to be tested together

It appears therefore that if parameters are found by comparing numbers of observations in ranges about the extreme values of the departure considered, each extending 2/3 of the way to the nearest point of agreement, there will be little loss of accuracy in comparison with the best possible solution, and the significance tests will not be greatly altered The change,

such as it is, will be represented by taking $2n/3$ for n in the sampling tests, while using the corrected standard errors

4—*Departures from a non-uniform distribution* Even if the suggested distribution is not uniform its determination is fundamentally a problem of sampling We could regard the observations as grouped by ranges of the argument and use their numbers to estimate the true chances of an observation lying in the respective ranges by the usual theory of sampling, the chances being taken as initially unknown except for their sum being unity The prior probability for this problem has been given (Jeffreys 1936, p 422) and leads to the maximum likelihood solution Laws of error connecting the chances are essentially subsidiary to the main problem of estimating the chances To test whether a given law of chance, for instance the normal law, agrees with an observed distribution, we cannot use the individual observations and apply the χ^2 test, since this test applied to individual observations will always appear to give agreement * The observations must be grouped in ranges before the test is applied, and the grouping at once introduces the conditions appropriate to sampling The fact that a change of the parameters in a law will in general introduce changes in the chances for different ranges that are not in general connected linearly implies that when the law is taken as part of the data we no longer regard the chances as initially unknown and unrelated, but when we wish to test divergences from a suggested law the problem of estimating the chances must be rediscussed as such

If the suggested law is that the chance for a range dx is $f(x)\,dx$ and we wish to discuss a modification $\{f(x)+ag(x)\}\,dx$ (the integral of $f(x)$ over the permitted range being unity and that of $g(x)$ zero), the chance in any finite range is linear in the unknown a Since in sampling problems we take the prior probability of a chance as uniformly distributed over the range permitted, that of a must be taken as uniformly distributed The point is that if we were to use only one range to estimate a we should estimate the chance of an observation in this range by simple sampling, taking its prior probability as uniformly distributed, and if $f(x)$ and $g(x)$ are given functions this implies that uniform distribution for the prior probability of a is the only form consistent with the rules that we have adopted already, and special rules can be adopted only for special reasons The limits permissible for a will then be given by the condition that $f(x) + ag(x)$ cannot become negative in the range permitted for x, and the rest of the work proceeds as before Thus no new hypothesis is needed

---

* If the grouping is so fine that not more than one observation can occur in any group, and the expectations in all groups, on the hypothesis to be tested, are equal, there may be m groups, n of which contain one observation and the rest none, χ^2 evaluated for this distribution will be $m-n$ whether the suggested law is true or not Nobody would use χ^2 for so fine a grouping

Here also grouping at wide intervals may be convenient We can subdivide the range and retain about two-thirds of the observations about the maxima of $g(x)$, the numbers in the ranges retained where $g(x)$ is positive and negative respectively can then be compared with the ratios predicted by the trial law $f(x)$ A significant departure will show that $f(x)$ is not the best form attainable on the data, and it will be possible to assert at least that the correct law implies modifications in the sense indicated by $g(x)$

SUMMARY

A method is developed for testing whether a series of observed frequencies supports a uniform distribution of chance or a suggested departure from it, and is applied to linear departures A modification applicable to grouped data is given, and a possible way of extending the method to test departures from suggested non-uniform laws of chance is described

REFERENCES

Fisher, R A 1936 "Statistical Methods for Research Workers" Edinburgh: Oliver and Boyd

Jeffreys, H 1936 *Proc Camb Phil Soc* **32**, 432–40

— 1937 *Proc Roy Soc* A, **162**, 479–95

New bands ending on the $1s\sigma\ 2p\sigma\ {}^1\Sigma_u$ state of H_2

By O W Richardson, F R S

Yarrow Research Professor of the Royal Society

(*Received* 8 *December* 1937)

In a recent paper (Richardson 1937) I referred to the importance and interest of the ^1X state of H_2 This is much deeper than any other state then known which goes down to the well-known odd state $1s\sigma\ 2p\sigma\ {}^1\Sigma_u$ on which the strongest lines in the visible part of the H_2 spectrum end In 1932 I decided that ^1X must have the configuration $1s\sigma\ 2s\sigma\ {}^1\Sigma_g$, a state which should give strong transitions down to $1s\sigma\ 2p\sigma\ {}^1\Sigma_u$ and which otherwise would unaccountably be missing I came to this conclusion in spite of some really quite strong arguments against it, which had been put forward by Weizel (1930), who considered ^1X to be a state with two electrons excited, to which he assigned the configuration $(2p\sigma)^2\, {}^1\Sigma_g$ An account of both these points of view may be found in my book (Richardson 1934, pp 307–12)

As a matter of fact the first suggestion that ^1X was $1s\sigma\ 2s\sigma\ {}^1\Sigma$ was first made tentatively by Dieke (1929), p 450, shortly after this state had been discovered It is very satisfactory to note that (Dieke 1936*a*), by a comparison of the spectra of H_2 and D_2 and by an extension of the spectrum further into the infra-red, using the recently developed photographic technique in that region, he has been able to prove that ^1X undoubtedly is $1s\sigma\ 2s\sigma\ {}^1\Sigma_g$

The main reasons why this very natural configuration for ^1X was not immediately acceptable were as follows In the first place until Dieke's (1936*a*) work was published the numbering of the vibrational levels was always doubtful In our account (Richardson and Davidson 1929, p 82) of the $^1X \rightarrow 2p\, {}^1\Sigma$ system which described the discovery of the ^1X state we definitely queried the tentative quantum numbering we assigned to all the initial vibrational (^1X) levels I have continued to query this numbering in every subsequent publication referring to the ^1X state In fact, the first account of the "Infra-red System" ($^1X \rightarrow 1s\sigma\ 2p\sigma\ {}^1\Sigma$) (p 81) opens with the words "It is probable that there is a great deal more of this system beyond the present limit of the infra-red measurements", and then goes on to justify this statement This situation had at least two important effects It made it possible that the depth of the electronic level of ^1X might be some thousands of wave numbers below that which was regarded as probable The repercussion of this particular uncertainty on the question of the identity of ${}^1X$ with $1s\sigma\,2s\sigma\,{}^1\Sigma$ was not so very severe at that time, as the identity of the experimental states which had been tentatively assigned to $1s\sigma\,3s\sigma\,{}^1\Sigma$ and $1s\sigma\,4s\sigma\,{}^1\Sigma$ was also insecure The second effect was the more formidable, for it made the value of the moment of inertia of the ${}^1X$ state of H_2 doubtful by possibly 30 % The value estimated with the tentative numbering was quite abnormal (too high) for $1s\sigma\,2s\sigma\,{}^1\Sigma$

The next difficulty was that the $1s\sigma\,2s\sigma\,{}^1\Sigma$ state was expected to have a very regular structure, whereas the explored part of the ${}^1X$ state was characterized by a structure of pronounced irregularity both rotational and vibrational There is no known cause of an irregularity of this type except a perturbation by some other state This led to another difficulty In molecular spectra, according to current theory, largely due to Kronig and now substantiated by a considerable body of experimental support, states capable of mutual perturbation are subject to a very considerable number of restrictions Among those important in the present case are (1) the energy levels of the states concerned must either approach close to or overlap one another, (2) the states must have values of the quantum number Λ differing by 0 or ± 1, the same multiplicity, not very different values of the equilibrium internuclear distance and the same type of symmetry, odd or even Now $1s\sigma\,2s\sigma\,{}^1\Sigma$ must be an even singlet state, and if it is to be identified with ${}^1X$ the state which perturbs it must also be, among other things, an even singlet state with a not too different energy The difficulty now arises that ${}^1X$ is a deep-lying state on the energy scale of H_2 states, and it is safe to say that there is no possible unknown normal singlet state of H_2, if ${}^1X$ is $1s\sigma\,2s\sigma\,{}^1\Sigma_g$, which could lie nearly so low as ${}^1X$ By normal I mean a state with only one of the two electrons in H_2 excited

States attributed to atoms with more than one excited electron are not common in spectroscopy A few have been found in atomic spectra, the best known examples occurring in the spectra of the alkaline earth metals, as, for example, calcium, while some of the corona lines have been attributed, but with how much certainty I am unable to judge, to transitions from doubly excited states of He In the case of molecular spectra the situation in this respect is not so clear, as the electronic configuration of so few excited states is known with certainty However, in the case of H_2 there are many states which are quite incomprehensible if only one electron is excited (Richardson 1934, pp 306–17, Richardson and Rymer 1934*a*, *b*)

The situation now is that we have to accept as a fact that, as a result of the work of Dieke (1936*a*), the state formerly called ${}^1X$ or, alternatively, the infra-red state, is actually $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ To account for the irregularities in the structure of $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, the existence of which is established beyond any possibility of doubt, Dieke (1936*a*) postulates the existence of an undiscovered state of H_2 lying in its neighbourhood and perturbing it If such a state exists and is able to perturb $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, even before we find it we have quite a lot of useful information about its properties In the first place it must be a state which lies low down on the energy scale of excited H_2 states, probably in the neighbourhood of $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ and $1s\sigma\, 2p\pi\, {}^1\Pi_u$ To perturb $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, its quantum number Λ must be connected with that of $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ by the relation $\Delta\Lambda = 0$ or ± 1 Since $\Lambda = 0$ for a Σ state, the postulated state must have $\Lambda = 0$ or 1 It must therefore be either a Σ or a Π state It must also have the same multiplicity and even or odd symmetry as $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ It is therefore either a ${}^1\Sigma_g$ or a ${}^1\Pi_g$ state It must also have a B_e and r_e (equilibrium internuclear distance) not very much different from the corresponding quantities for $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$

This information about the postulated states settles quite a lot about the emission bands to which it can give rise In the H_2 spectrum the rule which prohibits transitions between states of different multiplicity is universally obeyed so far as can be discovered, so that, this being a singlet state, it cannot emit bands ending on any triplet state In H_2 the $g \to g$ prohibition also is always respected The state therefore cannot give rise to bands ending on the ground state, $1s\, 1s\, {}^1\Sigma_g$, or the ${}^1X$ state, $1s\, 2s\, {}^1\Sigma_g$, or any other even singlet state In fact there are only two states, $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ and $1s\sigma\, 2p\pi\, {}^1\Pi_u$, down to which it could have transitions, all the other odd singlet states lying well above it We have seen that there are reasons for expecting its electronic energy to be close to that of $1s\sigma\, 2p\pi\, {}^1\Pi_u$, thus any transitions to this state would only involve small interchanges of energy and would not be expected to give rise to band lines of frequencies as high as the photographic infra-red limit $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ lies nearly 9000 wave numbers below $1s\sigma\, 2p\pi\, {}^1\Pi_u$, so that we have a reasonable expectation of bands ending on this state and with most of the intensity in the photographic infra-red In fact the expected location is not far from that of the $1s\sigma\, 2s\sigma\, {}^1\Sigma \to 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ bands, as the new state must be fairly near to $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ in order to perturb it

Although there is no great amount of unallotted strength in the part of the H_2 spectrum where this band system is expected, another set of circumstances tended to make its discovery not too unpromising All the known states, and there are a large number of them, which have transitions down to $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ give rise to band systems characterized by long progressions These are usually easier to detect than the band systems with practically all the strength along the diagonal axis which arise from transitions between states with values of r_e nearly equal to each other It is fortunate, too, that the vibrational, as also the rotational, differences of the final state $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ are known with great accuracy over a wide range of quantum numbers If we can find lines forming long progressions extending over several vibrational quantum numbers and continuing to satisfy the vibrational combination differences over a range of 4000 to 8000 wave numbers, that is to say 4 to 8 successive vibrational levels, with accuracy, we can be reasonably sure that we are dealing with a real band system and not with a succession of numerical coincidences This will be true, as I shall explain later, even if there are considerable gaps in the lines of such progressions

It was with this background and with a few other ideas less definitely formulated that I set out to look for this band system early this year I had not much difficulty in obtaining evidence of its existence My chief trouble has been that I found more than I expected It appears that there are two nearly coincident band systems ending on $1s\sigma\,2p\sigma\,{}^1\Sigma_g$ in this region It is true that one of them is weaker than the other and might possibly be a collection of coincidences, but I doubt it Another possibility is that there is really only one system and that what I have found is a jumble of this with a collection of coincidences This can easily happen in the interpretation of a complicated spectrum and has, in fact, already happened several times in connexion with the spectrum of H_2 The only way to make certain, that I know of, is by a comparison with the spectra of D_2 and of HD These are being measured, by others, but the results are not yet worked out and published Also this method might prove difficult, even possibly indecisive, so I have resolved to set out what I have found and the best I can make of it and to leave it to others to find out what, if anything, is wrong with it, if they can

As most of the suitable Roman capitals have been used up as provisional designations of excited H_2 states and it is undesirable to use any more or introduce a new system unless a large number of new states are found, an event which is quite likely to happen, I propose to call the upper states provisionally ${}^1\Sigma\,(a)$ and ${}^1\Sigma\,(b)$ Whatever else they may be, and this is a difficult point I discuss at length at the end of the paper, they are both clearly ${}^1\Sigma_g$ states They have transitions down to $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ giving rise to bands which have *P*, *R* but no *Q* lines and also have *P* 1 lines These facts are all that is required to establish that they are ${}^1\Sigma_g$ states

The band lines are set out in Tables I–VI In these tables the wave numbers of the band lines are followed by the experimental eye estimates in brackets of the relative intensities of the lines These are not very reliable guides to the relative intensities except locally in small regions of the spectrum The most extensive and probably most reliable are those of Gale, Monk and Lee (1928) In general all these eye estimates are relatively far too high in the red, yellow, green regions compared with the real energy values This is not a matter of small differences The eye estimates may be out by a factor approaching 100, possibly more The ideas of different observers about the meaning to be attached to a given set of intensity numbers in a given region are not very accordant either All this is important in connexion with the systems under consideration, which extend from the visible region to the present limit of the photographic infra-red It is certain from the work of Davidson (1932, p 593) that lines marked (1) by Gale, Monk and Lee (1928) near their infra-red limit may be as intense as lines they mark (10)—their highest intensity indicator—in the part of the spectrum centred about the yellow (around 18,000 wave numbers or 5500 A) There is thus much more strength in the infra-red lines of the present bands than the published eye estimates would appear to indicate

The letters after the intensity numbers in brackets have the following meanings The letter A stands for Allibone (1926) Preceding this and within the brackets *a* means a line observed on plate *A* only and *b* a line observed on plate *B* only D refers to Deodhar (1926) Both he and Tanaka (1925) use the letters *p*, *q*, *r* to indicate three successively weaker categories of faint lines all of which are less intense than Merton and Barratt's (1922) lowest intensity (0) Di refers to Dieke (1936 *a* or *b*) Most of these wave numbers are the result of a private communication and are preliminary values subject to revision and not yet published I have evidence, however, that those between 9750 and 8850 wave numbers are very accurate G refers to Gale, Monk and Lee (1928) Here *a*, *b*, *c*, *d* mean wave numbers of lines for which the measurements show an exceptional disagreement, in that order, *a* being closest to the normal error *h* means probably double Lines with any of these letters after them are probably blends M refers to Merton and Barratt (1922) and T to Tanaka (1925) (see also Deodhar, above) In Merton and Barratt's tables — means a line found only under exceptional circumstances and thus difficult to assign an intensity value to Tanaka also uses — to indicate a weak line whose intensity estimate he found baffling P stands for Poetker (1927) In his notation *h* means hazy, diffuse *d*, perhaps double *b*, broad *v*, shades towards violet, *r*, towards red * a line for which the measurements show exceptional disagreement All such lines are likely to be blends There is a systematic error in his measurements and the figures in the tables have had 0 30 wave number subtracted from his values in order to take most of this out

K refers to Kapuscinski and Eymers (1929) These are the only accurate intensity data based on a true energy scale which are available The values are given as numbers, e g 7 6, not enclosed in brackets The data do not extend into the infra-red but some idea of the real intensities of the infra-red lines can be got from the published eye estimates of Gale, Monk and Lee (1928) and of Poetker (1927) by means of the relations given by Davidson (1932)

The symbol † means claimed as a line in another band, ?† means claimed as a doubtful line in another band, †?† means claimed as an established line in another band and also as a doubtful line in a third band, and so on ??† means claimed as a very doubtful line elsewhere

If the bands are $\Sigma_g \rightarrow \Sigma_u$ bands, as we have been led to expect, there will be no Q branches on account of the selection rule $\Delta K = 0$ or ± 1 and the $s \leftarrow\!/\!\rightarrow a$ prohibition which rules out $\Delta K = 0$ in this case There should, however, be P branches ($\Delta K = +1$) and R branches ($\Delta K = -1$) In any v'' progression every $P(K+1)$ line and every $R(K-1)$ line will start from the same initial level, the level whose rotational energy would usually be represented by $F'_v(K)$ This means that in each such band the $R(K-1) - P(K+1)$ wave number differences are equal to the double rotational differences $F''(K+1) - F''(K-1)$ of the final state $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ at the corresponding v'' and, at the same time, the differences of the $R(K-1)$ and $P(K+1)$ lines at successive values of v'' are equal to the successive $\Delta v''$ differences of $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ at the same $K-1$ or $K+1$ as the case may be Thus when the wave numbers of any $R(K-1)$ or $P(K+1)$ line in a v'' progression is known the wave number of every single one of the remaining $R(K-1)$ and $P(K+1)$ lines in the progression is fixed by the known rotational and vibrational intervals of $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ As these differences are known with certainty and high accuracy from the work of Richardson and Davidson (1929) this restriction on the lines is a very severe one It is unlikely that many cases will be found of lines which satisfy it as a result of chance coincidences, at any rate for progressions extending to a considerable number of $R(K-1)$ or $P(K+1)$ lines

In order to set out in the tables the extent to which this requirement has been satisfied the following procedure has been adopted In each group of associated $R(K-1)$, $P(K+1)$ lines, in every progression, one line has been chosen for which the measurement is thought to be about correct and under the wave number of that line the letter O is placed Under each of the remaining associated $R(K-1)$, $P(K+1)$ lines will be found a wave number preceded by the letter d (defect) Thus d +0 08 means that the measured wave number, of the line, under which it is placed, is less by 0 08 (that is to

TABLE I—$^1\Sigma_g\,(a)$, $x \rightarrow 1s\sigma\; 2p\sigma\; ^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2
P 1	11172 54(1) P	—	—
R 0	—	9923 2 P	—
P 2	—	—	—
R 1	11255 56(1) P O	9939 9 P d −0 75	—
P 3	—	—	—
R 2	11261 17(1)† P d +0 18	—	8672 29(1)† O D_1 d −0 09
P 4	—	9694 93(00) D_1 O	—
R 3	11264 0(0) G	—	—
P 5	—	—	—
R 4	11262 9(05) G d −0 14	—	8697 49(1) D_1 O
P 6	—	—	—

TABLE I (*cont*)—$^1\Sigma_g\,(b)$, $z \rightarrow 1s\sigma\; 2p\sigma\; ^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2
P 1	—	—	?8805 57(0)† D_1 or (ca 8800 5)
R 0	11471 26(0*) P d +0 22	—	8872 09(00) D_1 d −0 45
P 2	—	—	8766 18(00) D_1 O
R 1	11488 41(1**h*) P O	—	8891 66(2)† D_1 d +0 50
P 3	—	—	—
R 2	11488 41(1**h*) P O	—	8898 48(0)† D_1 d +0 78
P 4	11222 25(0) P d −0 58	?9923 2 P d −1 21	8656 34(00) D_1 d +0 04
R 3	11471 26(0*) P	—	—
P 5	—	—	—
R 4	11447 94(1**h*) P d −0 17	—	—
P 6	—	—	8509 81(0)† D_1 O

TABLE II—$^1\Sigma_g\,(a),\ x+1 \rightarrow 1s\sigma\,2p\sigma\,{}^1\Sigma_u,\ v''$

$v'' \rightarrow$	0	1	2
P 1	13505 25(0) A d +0 43	12189 27(1) P O	10910 3(0) P d −0 87
R 0	13571 0(0*a*)† G O	12252 7(00) G d −0 04	—
P 2	13454 6(0) G d +0 25	12141 85(1)† P d +0 45	—
R 1	13582 6(1)†† G O	—	—
P 3	13390 0(0)† G d +0 24	?12083 8(8)† G d −0 50	—
R 2	13582 6(1)†† G O	—	—
P 4	13315 7(1*a*)† G d +0 16	12016 5(00*a*)?† G d −0 32	10749 8(0*h*) P d +0 79
R 3	13573 4(1)† G O	—	—
P 5	—	11944 1(1*a*) G d −0 33	—
R 4	13555 8(3)† G d −0 12	12256 0(00) G O	—
P 6	13148 3(4)† G d −0 32	11866 8(00)† G d +0 24	—

TABLE II (*cont*)—$^1\Sigma_g\,(b),\ z+1 \rightarrow 1s\sigma\,2p\sigma\,{}^1\Sigma_u,\ v''$

$v'' \rightarrow$	0	1	2	3
P 1	13600 4(4)?† G d 0 00	12284 00(0) P O	11004 3(1*h*) P d −0 15	—
R 0	—	—	—	—
P 2	13548 8(0)†?† G O	12236 21(0*) P d +0 04	—	??9720 3 P d −3 03
R 1	13677 0(5)† G d −0 34	—	11080 7(2*h*) P d −0 29	—
P 3	13484 3(1)† G O	12177 32(1) P d +0 04	—	—
R 2	13675 6(2) G d +0 05	12363 1(0)?† G O	11086 2(1*h*) P d +0 3	—
P 4	13408 9(1) G d +0 01	12109 4(1) G d −0 17	10844 1(1*bd*) P d −0 46	9611 00(00*d*) D1 d +0 03
R 3	13669 5(1)† G d −0 20	—	—	—
P 5	—	12039 67(1) P O	—	—
R 4	13663 3(0*a*)† ?† G d −0 10	12363 16(2)?† P d +0 36	—	—
P 6	13255 5(2*a*)† G O	—	—	—

TABLE III—$^1\Sigma_g\,(a)$, $x+2 \rightarrow 1s\sigma\, 2p\sigma\, ^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2
P 1	15723 84(0) G O	14407 43(1) G d 0 00	—
R 0	15787 23(5)† ?† ?† G d +0 50	14469 22(0) G d +0 17	—
P 2	15671 59(2)† G d −0 01	14359 08(2) G d −0 05	13082 4(1) G d +0 03
R 1	15800 76(10)† ?† G d −0 51	14483 84(0) G O	13203 6(1*a*)† G d +0 40
P 3	15607 91(0) G d −0 02	14300 97(1) G d −0 02	13029 29(1) A d −0 15
R 2	15802 72(0*a*) G d +0 07	14490 24(2) G O	—
P 4	15536 11(2) G d −0 06	14236 3(2*a*) G d +0 07	—
R 3	—	14491 21(1) G d +0 02	—
P 5	15459 91(0) M d −0 38	14168 84(1) A d −0 32	?12911 0(00) G d +0 50
R 4	15787 23(5)† ?† ?† G d +0 66	14488 24(00) ?† G d −0 03	—
P 6	?15380 80(0) G d −0 61	14099 2(0) G d +0 05	12850 38(1*h*) P d −0 45

TABLE III (*cont*)—$^1\Sigma_g\,(a)$, $x+2 \rightarrow 1s\sigma\, 2p\sigma\, ^1\Sigma_u$, v''

$v'' \rightarrow$	3	4	5
P 1	11882 3(0) G d +0 04	—	—
R 0	11941 2(1*c*) G O	10728 0(0) P d +0 19	—
P 2	11840 1(2*a*) G d −0 05	10631 4(0) P d −0 41	—
R 1	11958 8(0) G d −0 05	—	—
P 3	11791 0(0*c*) G d +0 02	10585 4(0) P d +0 45	—
R 2	11971 3(00*a*) G d −0 04	10762 9(1) P d −0 70	—
P 4	11738 1(0*b*)† G d +0 07	—	—
R 3	11981 3(0) G O	?10777 3(0) P d −1 27	—
P 5	11684 63(0*) P d −0 02	—	—
R 4	11990 01(1) P O	?10788 8(0) P d +1 11	9621 64(00) D$_1$ d +0 24
P 6	11631 94(1) P d −0 03	—	—

TABLE IV—$^1\Sigma_g(a)$, $x+3 \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2	3
P 1	17846 77(3) G 5 8K O	16530 23(5)† G d +0 13	—	14005 4(2)† G d −0 13
R 0	17902 12(0) G O	—	—	14055 6(3)† G d +0 01
P 2	17785 83(5)† G d +0 14	—	—	—
R 1	17906 43(2) G 6 7K O	—	—	14064 7(0)† G d +0 23
P 3	—	—	—	—
R 2	17903 73(0*a*) G O	—	15314 83(0*ah*) G d −0 25	14072 2(1)??† G d 0 00
P 4	—	—	—	—
R 3	17898 62(1)?† G d −0 05	—	—	14081 8(1)† G O
P 5	—	16268 73(*r*) D d +0 26	—	—
R 4	17889 06(0) G O	16589 34(0) G d +0 04	?15322 31 — M d +1 48	‡ ‡
P 6	17481 61(1)† G d −0 25	—	—	13733 0(1)† G d +0 08

‡ Lines absent in Gale, Monk and Lee's tables but might be part of an unresolved doublet in Allibone's

TABLE IV (*cont*)—$^1\Sigma_g(a)$, $x+3 \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, v''

$v'' \rightarrow$	4	5	6	7
P 1	—	—	—	—
R 0	?12843 07(0) P d −0 49	—	—	—
P 2	—	11569 39(0) P d −0 34	—	—
R 1	—	—	—	—
P 3	—	—	—	—
R 2	—	11687 09(0**h*) P d −0 28	—	9430 78(00) D1 d −0 01
P 4	—	11471 2(0c) G d −0 22	—	—
R 3	—	11704 0(00)† G d +0 11	—	—
P 5	—	11429 4(0*b*) G d −0 45	—	‡
R 4	12891 33(2) P d −0 25	11722 87(5) P d +0 18	??10585 4(0) P d +1 5	—
P 6	—	?11390 08(0*) P d +0 75	—	‡

‡ The interval leading to these two lines is unknown

TABLE V—$^1\Sigma_g\,(b)$, $z+2 \rightarrow 1s\sigma\,2p\sigma\,^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2
P 1	15733 48(rd^2) D O	—	13137 12(1)† P d +0 11
R 0	15795 05(2)† G O	—	13195 2(1)† ?† G d +0 03
P 2	—	—	—
R 1	15802 72(0*a*)† G d +0 34	—	—
P 3	15610 91(1)† ?† G d −0 55	—	—
R 2	15802 72(0*a*)† G d +0 07	14490 24(2) G O	—
P 4	15536 11(2) G d −0 06	14236 3(2*a*) G d +0 07	—
R 3	15794 25(1*a*) G O	—	?13214 9(00) G d +0 60
P 5	15455 31(*pd*) T d +0 30	—	?12907 21(1) P d −0 63
R 4	15779 68(0) G O	—	13214 58(1) A d −0 17
P 6	15371 61(*rd*) D d +0 36	?14090 4(4)† G d +0 63	—
Alternative or additional			
R 1	15821 82(1) G d +0 09	—	13225 5(0) G d +0 16
P 3	15629 18(0) M d +0 38	—	13050 8(1*a*) ?† G O

TABLE V (*cont*)—$^1\Sigma_g\,(b)$, $z+2 \rightarrow 1s\sigma\,2p\sigma\,^1\Sigma_u$, v''

$v'' \rightarrow$	3	4	5
P 1	11892 1(0*a*) G d −0 12	—	—
R 0	11948 6(3)† ?† G d −0 08	??10733 2(0) P d +2 31	—
P 2	—	—	—
R 1	—	10749 8(0*h*) P O	—
P 3	—	—	—
R 2	11971 3(00*a*) G d −0 04	10762 9(1) P d −0 70	—
P 4	11738 1(0*b*)† G d +0 07	—	—
R 3	—	—	—

TABLE V—(*cont*)

$v'' \rightarrow$	3	4	5
P 5	11681 19(1*bh*) P d −0 50	—	—
R 4	?11981 3(0) G d +0 50	—	9613 65(0) D_1 d +0 02
P 6	—	10437 1(0) P d −0 16	—
Alternative or additional			
R 1	—	—	?9588 2 P d +1 68
P 3	?11812 24(1) P d +0 46	—	9434 87(00) D_1 d −0 10

say it has a defect of +0 08) than that calculated from the wave number of the associated $R(K-1)$, $P(K+1)$ line under which the letter O is placed, when the appropriate cascade of combination differences is used These statements about associated $R(K-1)$, $P(K+1)$ lines apply equally to progressions of P 1 lines, for which there are no R lines with the same upper level, negative K values being inadmissible

In interpreting these defects a good deal of discrimination is necessary For pure lines of normal structure the defects should not be much in excess of ±0 10 if the measurements are those of Gale, Monk and Lee (1928) or of Dieke (1936*b*) (provided these are for wave numbers higher than 8840) Even when blends are involved I doubt whether there is a single authenticated case in Gale, Monk and Lee's tables with an error as high as 0 40 wave number Errors approaching this are not very rare in the tables of Merton and Barratt (1922), Allibone (1926) and Poetker (1927) Most of them arise from inferior resolution Lines with errors approaching 2 00 wave numbers may mean something in the tables of Tanaka (1925) and Deodhar (1926) These were for the most part made on plates taken by Merton under some special conditions, not particularly suitable for measurement but which seem to have brought up a number of lines not normally present Also in Poetker's (1927) tables there are a number of lines with wave numbers below about 11,000 as to the accuracy of which there is no known satisfactory check Several of these lie close to calculated values for band lines of the present systems Among them I have included a number with errors so large that they would be discarded if they occurred in the part of the tables with wave numbers above 11,000

Each system appears to be represented by four initial (upper) vibrational levels As the correct vibrational numbering is not certain I have called

TABLE VI—$^1\Sigma_g(b)$, $z+3 \rightarrow 1s\sigma\, 2p\sigma\, ^1\Sigma_u$, v''

$v'' \rightarrow$	0	1	2	3
P 1	—	16474 52(0) G d +0 10	15195 00(00)† G d −0 22	13949 53(1) A O
R 0	17850 87(7) G 23 6K O	16532 87(0*a*)†† G ? d −0 35	15253 57(4*h*) G d −1 52	—
P 2	17734 63(00) G d +0 09	?16420 14(*rd*) D d +2 03	—	—
R 1	17855 86(0)† G d +0 34	16539 82(00) G d −0 03	15260 81(*r*) T d −0 85	14014 7(1) G O
P 3	17663 26(0) G d +0 58	16356 44(0*h*)† G d +0 46	—	—
R 2	17850 87(7) G 23 6K d −0 33	?16538 98(2)† G d −0 99	15260 81(*r*) T d +0 59	—
P 4	17583 80(0) G O	16283 83(2)† G d +0 39	—	?13787 17(6) A d −1 15
R 3	17837 29(3)† G O	16530 23(5)† G d +0 12	—	—
P 5	17498 65(2)† G d 0 00	—	—	?13724 74(1) A d −1 01
R 4	—	?16522 24(*rd*) D d −1 70	?15253 57(4*h*) G d +1 38	14022 33(0*b*) A d +0 01
P 6	17412 52(1*a*) G O	16132 09(5)† G d −0 49	14881 60(2)† A d +0 66	13663 3(0*a*)†?† G d +0 94

TABLE VI (*cont*)—$^1\Sigma_g(b)$, $z+3 \rightarrow 1s\sigma\, 2p\sigma\, ^1\Sigma_u$, v''

$v'' \rightarrow$	4	5	6	7
P 1	—	—	—	—
R 0	12790 53(0*d*, *h*) P d +0 80	—	10464 6(1*h*) P d −0 72	9348 80(1)† D_1 d −0 42
P 2	—	—	—	?9262 81(00*d*) D_1 d −1 04
R 1	12801 72(1*h* *r*) P d +1 22	—	—	—
P 3	—	—	—	—
R 2	12809 9(1*a*)?† G d +0 05	—	—	—
P 4	?12586 8(1*b*)?† G d −0 88	—	—	?9123 14(00*d*) D_1 d +0 61
R 3	12816 05(0*) P d −0 80	—	—	—
P 5	—	?11367 1(1*a*)† G d +0 52	—	9070 55(0) D_1 d −0 07
R 4	12822 4(0)† G d −0 16	—	—	—
P 6	—	—	—	—

these provisionally $v' = x, x+1, x+2$ and $x+3$ in succession for ${}^1\Sigma_g(a)$ and $v' = z, z+1, z+2, z+3$ in succession for ${}^1\Sigma_g(b)$ In each case I think there is little doubt that the successive v's increase by unity I also think it probable that x and z are each zero

It will be noticed that eight successive vibrational levels of the final state $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ are represented by something in the progressions that come from the $v' = x+3$ and $v' = z+3$ levels The two progressions each extend from the yellow to the present photographic infra red limit at about 8500 wave numbers These are long progressions even when compared with the progressions ending on $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ which contain the most intense lines known in the H_2 spectrum (there may possibly be more intense lines in the Lyman region, but the energy of these has not been measured) It is possible, even likely, that there is more of both systems further out in the infra-red It is unlikely to be a coincidence that each of the eight progressions terminates at the present infra-red limit If we denote by v''_m the highest value of v'' to which transitions have been found, the lines of the bands which go down to $v''_m + 1$ would, in each case, lie just beyond the present photographic infra-red limit

This puts a restriction on what can be asserted with confidence about the distribution of intensity in the bands, but that it is not a normal one is quite certain In the bands coming from typical states, such as $1s\sigma\,3d\sigma\,{}^1\Sigma_g$, down to $1s\sigma\,2p\sigma\,{}^1\Sigma_u$, the intensity distribution is of the usual Condon type for states having widely different values of r_e and B_e These have a minimum along the bands on or near the diagonal axis of the v', v'' diagram and maxima lying on a parabola which is symmetrical about this diagonal axis Now if we turn to Table IV we find eight lines in the band ending on $v'' = 0$, three in the band ending on $v'' = 1$, at most two at $v'' = 2$, at least six at $v'' = 3$, at most two at $v'' = 4$, about six at $v'' = 5$, a very doubtful one at $v'' = 6$ and at least one that looks good at $v'' = 7$ Thus in this progression there are at least three and possibly four values of v'' for which the intensity has a maximum and at least two and possibly three intervening values for which it has a minimum Turning to Table I in ${}^1\Sigma_g(b)$, $z \to 1s\sigma\,2p\sigma\,{}^1\Sigma_u$, v'' there is something to indicate that six lines are represented in the band at $v'' = 0$, there is only one line at $v'' = 1$, and that very doubtful, whereas there are about six lines at $v'' = 2$ The intensity distribution in ${}^1\Sigma_g(a)$, $z \to v''$ also appears to be similar to that in ${}^1\Sigma_g(b)$, $z \to v''$ If these were normal $v' = 0$ progressions the intensity of the bands should fall steadily from $v'' = 0$ to $v'' = 2$ The same kind of behaviour is shown also by the other progressions, but owing to the occurrence of blends and probably also perturbations it is not so distinct

Intensity distributions of this type are no new phenomenon in the H_2 spectrum Turning to Richardson (1934), on p 141 we have a picture of a normal intensity distribution, that for $1s\sigma\, 3s\sigma\, {}^1\Sigma_g \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, in fig 10, and on p 137, in fig 9, one of a distribution, for ${}^1\mathrm{M} \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, much resembling those of the systems I am describing On pp 126–32 will be found the lines of five systems which have intensity distributions of this type, some of them being among the strongest systems in the H_2 spectrum I believe these systems come from upper states which have both electrons excited, and I regard this similarity as support for attributing a condition of double excitation to the upper states of the present systems I have shown already that there are no known rational grounds for believing that an undiscovered state with a singly excited electron could exist anywhere near where these states lie

Theoretically, when the intensities of the lines are reduced by the a priori probability quantum number K factor and the Boltzmann factor the lines with even K'' should be three times as intense as those with odd K'' Actually this means that the lines with even K'' should have roughly three times the intensity of the average of the $K''-1$ and $K''+1$ lines of the same band if K'' is small When allowance is made for blends it can be seen from the tables that this condition is satisfied, although, owing to the various blends and perturbations and the difficulty of assessing and correlating the eye estimates of intensity, particularly when made by different observers, the agreement can only be regarded as qualitative for these band systems There are only three lines in the whole of the two systems for which the real intensities have been measured The line 17850 87(7) G 23 6 K which does duty as $R\,0$ and $R\,2$ of ${}^1\Sigma_g\,(b)\ z+3 \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, $v''=0$ is one of the two or three strongest lines of the H_2 spectrum hitherto unclassified, whose real intensities have been measured by Kapuscinski and Eymers (1929)

I have left in the tables a number of lines for which the combinational errors are so large that it is not very probable that these lines are really lines of the bands I have done this deliberately, with a view to saving time and trouble for those who are in a position to compare the spectrum of H_2 with those of D_2 and HD They should have little difficulty in deciding, in most cases, which, if any, of these doubtful lines are genuine band lines

Prior to the recent work of Dieke only nine lines with wave numbers less than 9750 were known These were measured by Poetker (1927) Dieke's (1936*b*) extension of the photographic infra-red limit to 8500 wave numbers has raised the number of lines beyond 9750 from 9 to 114 Out of these 114 lines, 39 are claimed by the systems $1s\sigma\, 3p\sigma\, {}^3\Sigma_u \rightarrow 1s\sigma\, 2s\sigma\, {}^3\Sigma_g$ and $1s\sigma\, 2s\sigma\, {}^1\Sigma_g \rightarrow 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ and 21 by the bands of the two systems here described Among these 60 lines, 6 are claimed twice and several are very doubtful In any event almost half the new lines are accounted for by four systems whose existence was established without knowledge of any of them

The $v' = x+1$ progression of $^1\Sigma_g(a)$ is not so well authenticated as most of them In fact it is not easy to be sure of this one as there are alternatives for most of the sets of lines attributed to it in Table II These are $P\,1$ lines $v'' = 0$, 13495 55(1*b*) A ab G, O, $v'' = 1$, 12179 3(00*a*)?† G, *d* −0 16, $v'' = 2$, 10899 0(0) P, d +0 30 $R\,0$ lines $v'' = 0$, 13555 8(3)† G, d −0 09, $v'' = 3$, 9709 18(00) D1, O $P\,2$ lines $v'' = 0$, 13439 5(00)† G, d +0 06, $v'' = 2$, ?10851 7(0*h*) P, d −1 29 $R\,2$ lines $v'' = 0$, 13568 4(0*a*)† G, O, $v'' = 1$, 12256 0(00) G, d −0 15 $P\,4$ lines absent $R\,3$ lines $v'' = 0$, 13571 0(0*a*)† G, O $P\,5$ lines $v'' = 1$, 11941 2(1*c*) G, 11941 57(2*h*) P, d of mean −0 015 $R\,4$ lines $v'' = 0$, 13561 8(2) G, O $P\,6$ lines $v'' = 2$, ?10623 1(1) P, d +0 74 If any of these sets of lines are substituted for those in Table II there would appear to be a perturbation at the $v' = x+1$ level

ROTATIONAL STRUCTURE

The rotational structure, or ensemble of rotational quantum level energy differences, of the lower state $1s\sigma\,2p\sigma\,^1\Sigma_u$ is well known and probably with more accuracy than that of any other state of any molecule, so that we do not need to describe it The level differences for the upper states can be got from the relations

$$\begin{aligned} F'(K)-F'(K-1) &= P(K+1)-P(K)+F''(K+1)-F''(K) \\ &= R(K-1)-R(K-2)+F''(K-1)-F''(K-2) \end{aligned}$$

The mean values of these intervals $F'(K)-F'(K-1)$ got from the appropriate lines of the various bands and the accurately known lower intervals $F''(K+1)-F''(K)$ give the structures of the upper states of the systems shown in Table VII

The $R\,1$, $P\,3$ lines which give the more regular structure for $^1\Sigma_g(b)$ at $v' = z+2$ are not very securely established There are only three lines, all of which appear to be blends, and the combinations are poor Also even with these lines the structure is irregular If the alternative $R\,1$, $P\,3$ lines at the bottom of Table V are substituted, the very irregular alternative structure shown at the bottom of Table VII is obtained It is possible that the level is split and that both sets of lines belong to the bands In any event there seems to be a perturbation here

TABLE VII—ROTATIONAL STRUCTURE OF $^1\Sigma_g\,(a)$

$v' \rightarrow$ $K \downarrow$	x		$x+1$		$x+2$		$x+3$	
5 *a*	—		—		—		—	
	150 58		134 08		141 42		142 25	
4 *s*	→	32 69	→	28 22	→	31 00	→	32 30
	117 89		105 86		110 42		109 95	
3 *a*	→	34 98	→	28 56	→	30 61	→	34 74
	82 91		77 30		79 81		74 61	
2 *s*	→	30 04	→	26 85	→	28 37	→	31 34
	52 87		50 45		51 44		43 27	
1 *a*	→	22 72	→	23 80	→	26 35	→	27 11
	30 15		26 65		25 09		16 16	
0 *s*								

ROTATIONAL STRUCTURE OF $^1\Sigma_g\,(b)$

$v' \rightarrow$ $K \downarrow$	z		$z+1$		$z+2$		$z+3$	
5 *a*	—		—		—		—	
	128 1		145 43		137 11		134 61	
4 *s*	→	30 2	→	36 52	→	30 59	→	33 13
	97 9		108 93		106 52		101 48	
3 *a*	→	20 6	→	32 73	→	29 49	→	30 17
	77 3		76 2		77 03		71 31	
2 *s*	→	21 4	→	25 5	→	30 17	→	27 47
	55 9		50 7		46 86		43 84	
1 *a*	→	25 0?	→	25 02	→	24 14	→	22 94
	?30 9		25 68		22 72		20 90	
0 *s*								

Alternative in $z+2$

5 *a*	—	
	137 11	
4 *s*	→	30 59
	106 52	
3 *a*	→	45 34
	61 18	
2 *s*	→	−1 53
	62 71	
1 *a*	→	39 99
	22 72	
0 *s*		

Apart from a few such irregularities, mainly due to perturbations, and to which I shall return in the next section, the rotational structure of all the bands is similar The second differences (shown in Table VII) which measure the effective value of the quantity B_K are smallest at $K = 0$ and increase at first and have a tendency to diminish or, at any rate, to become steady at the highest K value found ($K = 5$) This is an indication of uncoupling, and the uncoupled state has a smaller internuclear distance than the coupled state Phenomena of this kind are common in states which go down to $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ As I have remarked, the rotational structures of the two upper states I am now describing are much alike When compared with the others they show most resemblance to $3\,{}^1$K, ^{1}L, ^{1}M and ^{1}N, the resemblance to ^{1}M being particularly close Each of these four states has been assigned to some state in which both electrons are excited, the main ground for this conclusion being that there does not seem to be any possible molecular state with only one electron excited to which they could be attributed These resemblances are additional reasons for supposing that the two new upper states are also states with two excited electrons

On account of the irregularities in the structure of the bands it is not possible to give exact values of the rotational constants If these irregularities are all due to vibrational perturbations the approximate effective values are for ${}^1\Sigma_g(a)$, B_x = ca 15 1 wave numbers and r_x about $1\ 5 \times 10^{-8}$ cm , for ${}^1\Sigma_g(b)$, B_z = ca 13 6 wave numbers and r_z about $1\ 5_5 \times 10^{-8}$ cm If, as is not unlikely, there is also some rotational uncoupling, the values of B_x and B_z for K small would be somewhat less and those for K large somewhat higher than these values The other values are roughly $B_{x+1} = 13\ 7_5$, $B_{x+2} = 14\ 2$, $B_{x+3} = 13\ 4_5$, $B_{z+1} = 14\ 1_5$, $B_{z+2} = 13\ 6$ and $B_{z+3} = 12\ 9_5$ The values of the "constant" α in $B_n = B_0 - n\alpha$ are for $\Sigma(a)$, α = about $0\ 5_5$, and for $\Sigma(b)$ probably a little smaller

PERTURBATIONS

Some of these have been referred to already Returning to the $v' = z+2$ level of ${}^1\Sigma_g(b)$ this may have a structure even more perturbed than that indicated at the bottom of Table VII Referring to Tables III and V we find that the $K' = 3$ level ($R\,2$, $P\,4$ lines) is identical with the $K' = 3$ level of ${}^1\Sigma_g(a)$ at $v' = x+2$ The distribution of intensity in this progression makes it certain that it belongs to ${}^1\Sigma_g(a)$, whether it also belongs to the other state is not at all certain An argument in favour of retaining it in both systems is that there are two lines in it marked *a* which are not claimed elsewhere On the other hand it is quite possible that this level in the ${}^1\Sigma_g(b)$ system is displaced and quite weak, and thus difficult to find This might make the rotational structure very irregular

That this $\Sigma_g(b)$ level is perturbed is not surprising, because these two states ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ cross one another within the $v' = x+2$ levels for ${}^1\Sigma_g(a)$ and the $v' = z+2$ interval for ${}^1\Sigma_g(b)$ The actual point at which the two perturbed levels cross is very close to $K' = 3$ in each case Owing to some complications, which include the possible rotational uncoupling, I have not been able to decide exactly where the unperturbed levels cross There is some evidence of the mutual influence of this perturbation on the state ${}^1\Sigma_g(a)$, but it is not very marked This state ${}^1\Sigma_g(a)$ appears to be more stable than the other There is other evidence which supports this contention

Another level which has an irregular rotational structure pointing to a perturbation is ${}^1\Sigma_g(b)$ at $v' = z$ When the spacing of the rotational levels here is compared with a more normal structure spread over the same total energy interval it is found that the $K' = 0$, 1 and 2 levels lie above the normal values and the $K' = 3$, 4 and 5 levels below them The perturbing level therefore appears to lie between $K' = 2$ and $K' = 3$ of $v' = z$ of this state There can be little doubt that the perturbing state is $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, since the rotational energy levels of the $v' = 1$ level of $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ cross the rotational levels of ${}^1\Sigma_g(b)$ at $v' = z$ The perturbed levels of the two states cross one another between $K' = 2$ and $K' = 3$ It is very probable that the unperturbed levels also cross between $K' = 2$ and $K' = 3$, as the perturbation is small compared to the level spacing Dieke (1936*a*) remarks that there is a perturbation in $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ at $v' = 1$ for small K Possibly this is the counterpart of the one I have been describing The $v' = x$ level of $\Sigma(a)$, though rather close to $v' = z$ of $\Sigma(b)$, lies wholly below $v' = 1$ of $2s\, {}^1\Sigma$, so that no corresponding perturbation should arise and none is observed

We have seen (p 331) that the $v' = x+1$ progression of ${}^1\Sigma_g(a)$ is very probably affected by a perturbation The state causing this perturbation is also most likely to be $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ The $v' \to 0''$ lines of ${}^1\Sigma_g(a) \to 1s\sigma\, 2s\sigma\, {}^1\Sigma_u$ at $v' = x$ lie below the $v' \to 0''$ lines of $1s\sigma\, 2s\sigma\, {}^1\Sigma_g \to 1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ at $v' = 1$ at low values of K, whereas, the reverse is the case when we compare the corresponding lines of the successive pairs $(x+1)' \to 0''$, $2' \to 0''$ $(x+2)' \to 0''$, $3' \to 0''$ $(x+3)' \to 0''$, $4' \to 0''$ Thus the two states ${}^1\Sigma_g(a)$ and $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ cross between $v' = x$ and $v' = x+1$ for the first and between $v' = 1$ and $v' = 2$ for the second They should also recross within the limits of the $(x+1)' \to 0''$ and the $2' \to 0''$ bands for rather high values of K', in the neighbourhood of $K' = 6$, but lines with values of K' as high as this are unknown Evidently the conditions for a perturbation to occur in this neighbourhood are satisfied

If, as appears to be the case, all three mutually perturbing levels are Σ levels, the perturbations must all be due to interaction between electronic and vibrational energy according to a criterion mentioned by Dieke (1936*a*), since they all have the same value of Λ, namely zero

I do not wish to insist too much on these attempts to explain the observed irregularities The results do not appear to harmonize completely in all details with the properties of the state perturbing $1s\sigma\,2s\sigma\,{}^1\Sigma_g$, which were predicted by Dieke (1936*a*) from his study of the perturbations of $1s\sigma\,2s\sigma\,{}^1\Sigma_g$ However, there are so many possible causes of perturbation that the effects are rather likely to be mixed up and difficult to disentangle It does not seem worth while to push further in this direction until someone has examined the corresponding band systems of HD and D_2

THE VIBRATIONAL STRUCTURE OF THE UPPER STATES

The vibrational intervals have been obtained by subtracting pairs of corresponding lines of successive progressions, such as $R(K)$ of $(x+3)' \rightarrow v''$ and $R(K)$ of $(x+2)' \rightarrow v''$, $v'' = 0, 1, 2,$ etc Since $P(K+2)$ lines give the

TABLE VIII—VIBRATIONAL INTERVALS OF ${}^1\Sigma_g\,(a)$

$\Delta v' \rightarrow$	$(x+1)-x$		$(x+2)-(x+1)$		$(x+3)-(x+2)$
$\Delta_1 v'$ $(K=0)$	2332 9		2218 40		2122 98
$\Delta_2 v'$ $(K=0)$		114 5		95 2	
$\Delta_1 v'$ $(K=1)$ from $R\,0$	2329 48		2216 62		2114 50
$\Delta_1 v'$ $(K=1)$ from $P\,2$	2329 34		2216 74		2114 42
$\Delta_1 v'$ $(K=1)$ mean	2329 42		2216 64		2114 46
$\Delta_2 v'$ $(K=1)$ mean		112 78		102 18	
$\Delta_1 v'$ $(K=2)$ from $R\,1$	2327 0		2217 57		2106 19
$\Delta_1 v'$ $(K=2)$ from $P\,3$	2327 0		2217 63		2106 22
$\Delta_1 v'$ $(K=2)$ mean	2327 0		2217 60		2106 20
$\Delta_2 v'$ $(K=2)$ mean		109 4_0		111 40	
$\Delta_1 v'$ $(K=3)$ from $R\,2$	2321 2		2220 18		2101 12
$\Delta_1 v'$ $(K=3)$ from $P\,4$	2321 4		2220 13		2101 12
$\Delta_1 v'$ $(K=3)$ mean	2321 3		2220 16		2101 12
$\Delta_2 v'$ $(K=3)$ mean		101 1_4		119 04	
$\Delta_1 v'$ $(K=4)$ from $R\,3$	2309 4		2224 7_6		2100 50
$\Delta_1 v'$ $(K=4)$ from $P\,5$	2309 5_5		2224 9_9		2100 29
$\Delta_1 v'$ $(K=4)$ mean	2309 5		2224 8		2100 45
$\Delta_2 v'$ $(K=4)$ mean		84 7		124 3_5	
$\Delta_1 v'$ $(K=5)$ from $R\,4$	2292 9_1		2232 14		2101 16
$\Delta_1 v'$ $(K=5)$ from $P\,6$	2292 8_8		2232 2_4		2101 14
$\Delta_1 v'$ $(K=5)$ mean	2292 9		2232 17		2101 15
$\Delta_2 v'$ $(K=5)$ mean		60 7		131 02	

same values as $R(K)$ lines a considerable number of independent determinations can usually be made in this way The values for ${}^1\Sigma_g(a)$ are set out in Table VIII In this table the values got from the R lines and the P lines are given separately as well as the weighted means, so as to give some idea of the consistency of the different determinations The values for ${}^1\Sigma_g(b)$ are given in Table IX The alternatives in this table arise from the alternative $K' = 2$ lines given at the bottom of Table V I believe that the claims of these lines to belong to the bands are stronger than those in the body of the table, which give a more regular structure Even if the lines which give this more regular structure are adopted it will be seen from Table IX that there are some indications of a perturbation at or near $K' = 2$, $v' = z+2$ of

TABLE IX—VIBRATIONAL INTERVALS OF ${}^1\Sigma_g(b)$

$\Delta v' \rightarrow$	$(z+1)-z$		$(z+2)-(z+1)$		$(z+3)-(z+2)$
$\Delta_1 v'$ $(K=0)$ mean	?2198 63		2133 06		2057 52
$\Delta_2 v'$ $(K=0)$ mean		?64 57		75 54	
$\Delta_1 v'$ $(K=1)$ mean	2193 64		2130 10		2056 $_{01}$
$\Delta_2 v'$ $(K=1)$ mean		63 54		70 $_{09}$	
$\Delta_1 v'$ $(K=2)$ mean	2188 3_2		2126 07		2053 17
$\Delta_2 v'$ $(K=2)$ mean		62 2_5		72 90	
$\Delta_1 v'$ $(K=3)$ mean	2187 3		2127 16		2048 16
$\Delta_2 v'$ $(K=3)$ mean		60 1_4		79 00	
$\Delta_1 v'$ $(K=4)$ mean	2198 20		2124 62		2043 34
$\Delta_2 v'$ $(K=4)$ mean		73 58		81 28	
$\Delta_1 v'$ $(K=5)$ mean	2215 32		2116 5_1		2040 51
$\Delta_2 v'$ $(K=5)$ mean		98 8_1		76 0_0	
Alternative $K = 2$					
$\Delta_1 v'$ $(K=2)$ mean	2188 32		2144 84		2034 07
$\Delta_2 v'$ $(K=2)$ mean		43 48		110 77	

Arranged as Table IV of Richardson (1937, p 500) these would be

$K \rightarrow$	0		1		2		3		4		5
Mean $(z+1)-z$, R, P	?2198 63		2193 64		2188 3_2		2187 3		2198 20		2215 32
Δ		4 99		5 32		1 02		-10 9		-17 2	
Mean $(z+2)-(z+1)$, R, P	2133 06		2130 10		2126 07		2127 16		2124 62		2166 5
Δ		2 96		4 03		-1 09		2 44		8 12	
Mean $(z+3)-(z+2)$, R, P	2057 52		2056 01		2053 17		2048 16		2043 34		2040 51
Δ		1 51		2 84		5 01		4 82		2 83	

With alternatives corresponding to the alternative at $K = 2$

Mean $(z+2)-(z+1)$, R, P	2133 06		2130 10		2144 84		2127 16		2124 62		2116 5
Δ		2 96		-14 74		27 68		2 44		2 83	
Mean $(z+3)-(z+2)$, R, P	2057 52		2056 01		2034 07		2046 18		2043 34		2040 51
Δ		1 51		21 92		-12 11		2 44		2 83	

${}^1\Sigma_g\,(b)$, although it is very mild compared with the one obtained if the lines which I regard as more probable are adopted

That the detailed structure of ${}^1\Sigma_g\,(a)$ is highly irregular, or, at any rate, unusual, becomes evident if we arrange the vibrational differences of this state in the same manner as those of ${}^1\Sigma_g\,(b)$ at the bottom of Table IX This is done in Table X

TABLE X—VIBRATIONAL INTERVALS OF ${}^1\Sigma_g\,(a)$

$K \rightarrow$	0		1		2		3		4		5
Mean $(x+1)-(x)$, R, P	2332 9		2329 42		2327 0		2321 3		2309 5		2292 9
Δ		3 48		2 42		5 7		11 8		16 6	
Mean $(x+2)-(x+1)$, R, P	2218 40		2216 64		2217 60		2220 16		2224 8		2232 17
Δ		1 76		-0 96		-2 56		-4 64		-7 9	
Mean $(x+3)-(x+2)$, R, P	2122 98		2114 46		2106 20		2101 12		2100 15		2101 15
Δ		8 52		8 26		5 08		0 67		-0 70	

On account of these complications it is not possible to assign exact values to the vibrational constants of either of these two ${}^1\Sigma_g$ states For ${}^1\Sigma_g\,(a)$ the effective values appear to be close to $\omega_0 = 2391\,_3$ and $x\omega_0 = 59\,_6$ For ${}^1\Sigma_g\,(b)$ they are in the neighbourhood of $\omega_0 = 2233$ and $x\omega_0 = 35$

THE ELECTRONIC ENERGY OF THE UPPER STATES

The ν_0 of the $v = x$ level of ${}^1\Sigma_g\,(a)$ is close to 23154, and the ν_0 of the $v = z$ level of ${}^1\Sigma_g\,(b)$ is close to 22917, wave numbers below the ground state of the molecular ion $H_2{}^+$ If x and z are each equal to zero then these ν_0's are the ν_e's of the respective states and ${}^1\Sigma_g\,(a)$ lies about 2189 wave numbers above $1s\sigma\,2s\sigma\,{}^1\Sigma_g$, which lies about 25343 wave numbers below the ground state of $H_2{}^+$ Thus ${}^1\Sigma_g\,(a)$ has a denominator (z in $\nu_e = R_H/z^2$, $R_H =$ Rydberg constant for H) of about $2\ 10_0$ The corresponding quantities for ${}^1\Sigma_g\,(b)$ are 2426 wave numbers and $2\ 19_3$ The wave numbers are more accurate than these denominators, which have been got from them with a slide rule It is likely that x and z are both zero, but it is not certain, as any progressions starting further down in the infra-red would be difficult to establish owing to the paucity of measured lines

THE PROBABLE CONFIGURATION OF THE UPPER STATES

If these band systems are genuine there is only one possible or, at any rate, reasonable interpretation of the general character of their upper states It is that they are states with both the electrons in H_2 excited There are many reasons for thinking this, but there is among them one which to my mind would carry conviction if it stood alone The only states of H_2 with only one electron excited with its principal quantum number n less than 5 which can go down to $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ are, on account of various prohibitions all of which hold good in the H_2 spectrum, the following $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, $1s\sigma\, 3s\sigma\, {}^1\Sigma_g$, $1s\sigma\, 4s\sigma\, {}^1\Sigma_g$, $1s\sigma\, 3d\sigma\, {}^1\Sigma_g$, $1s\sigma\, 3d\pi\, {}^1\Pi_{ab}$, $1s\sigma\, 3d\delta\, {}^1\Delta_{ab}$, $1s\sigma\, 4d\sigma\, {}^1\Sigma_g$, $1s\sigma\, 4d\pi\, {}^1\Pi_{ab}$, $1s\sigma\, 4d\delta\, {}^1\Delta_{ab}$ All these are well known and their structure has been analysed $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$ has a Rydberg denominator close to 2, all the states with $n = 3$ (for the excited electron) have denominators close to 3 and all those with $n = 4$ have denominators close to 4 In addition there are several known states with $n = 5$ and all of them have denominators close to 5 If there were one state, much less two, with one electron only excited, and to $n = 5$ and having a denominator close to 2, it would be something in the nature of a miracle

In any event there are several other weighty arguments which point in the same direction Every band system, ending on $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ and coming from an upper state with only one electron excited, which is extensive enough for the intensity distribution among its constituent bands to be ascertained, has a comparatively simple type of intensity distribution In a v', v'' diagram the bands of maximum intensity lie on or close to a single Condon parabola This parabola is almost symmetrical about the line joining the bands of minimum intensity which is roughly coincident with the diagonal axis As I have shown (p 329), the distribution in the present system is much more complex and has a close resemblance to the intensity distributions in the systems ending on $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ which come from the upper states $3\,{}^1K$, ${}^1L$, ${}^1M$, ${}^1N$ and ${}^1Q$, and including also the system called the $\lambda 4142\ 8$ progression I have attributed the upper states of all these systems to states with both electrons excited on various grounds, the strongest being that there seems to be no other rational way of accounting for them

Again, the rotational structures of the two upper ${}^1\Sigma_g$ states are peculiar and similar to one another and they are also very like the unusual structures of the ${}^1\Sigma_g$ states $3\,{}^1K$, ${}^1L$, ${}^1M$ and ${}^1N$ These unusual but similar rotational structures suggest that all these six ${}^1\Sigma_g$ states have some unusual features in common I suggest that this common feature is that they all have both electrons excited

Again, the whole of the argument put forward by Weizel (1930) in order to identify the ${}^1X$ state with the doubly excited state $2p\sigma\, 2p\sigma\, {}^1\Sigma_g$ can be used equally well to identify either of the two new upper states with $2p\sigma\, 2p\sigma\, {}^1\Sigma_g$ The essential result of that argument was that $2p\sigma\, 2p\sigma\, {}^1\Sigma_g$ should be a ${}^1\Sigma_g$ state in the same neighbourhood as ${}^1X$ (now identified with

$1s\sigma\, 2s\sigma\,{}^1\Sigma_g$) As both the two new upper states are ${}^1\Sigma_g$ states and are very close to ${}^1X$, the argument will apply to either of them In my book (Richardson 1934, p 309) I criticized Weizel's argument not on the ground that it was wrong in principle but that it was directed towards a wrong objective The situation was rather complex, but on the whole I thought the arguments in favour of ${}^1X$ being $1s\sigma\, 2s\sigma\,{}^1\Sigma_g$ definitely outweighed those in favour of its being a doubly excited state, and this view has turned out to be correct

There is not a great deal that we can say about doubly excited molecular states from the theoretical point of view The energies and structures of none of them have been calculated But we can say something The states we are concerned with are all singlet states so, not to take up too much space, I shall confine my remarks to them According to Pauli's exclusion principle, which applies to all the part of the H_2 spectrum which is understood, two of the eight quantum numbers (n, l, λ, s) (n, l, λ, s) which define the state of the two electrons must be different In the singlets the spin quantum numbers have different values $+\frac{1}{2}$ and $-\frac{1}{2}$ The remaining quantum numbers n, l, λ can therefore be the same for each electron In the triplets the spin wave functions are symmetric, so that at least one of the other three pairs, n, n, l, l or λ, λ, must be different This makes the number of possible doubly excited states greater in the singlets than in the triplets

The deepest states, such as $2p\sigma\, 2p\pi\,{}^1\Pi_g$, of a set which differs only in the principal quantum number of one of the excited states, will be the first term of a potential Rydberg series, such as $2p\sigma\, 2p\pi\,{}^1\Pi_g$, $2p\sigma\, 3p\pi\,{}^1\Pi_g$, $2p\sigma\, 4p\pi\,{}^1\Pi_g$, $2p\sigma\, np\pi\,{}^1\Pi_g$, , which will terminate at $n = \infty$ with an excited state of the ion H_2^+, in this case the state $2p\sigma\,{}^2\Sigma_u$ Teller (1930) has worked out the energies of the deeper excited states of the ion as a function of the internuclear distance It is clear either from Weizel's argument or from my modification of it (Richardson 1934) that Teller's energy curve for the ion $2p\sigma\,{}^2\Sigma_u$ would admit of $2p\sigma\, 2p\sigma\,{}^1\Sigma_g$ being in the neighbourhood of either ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ or of $3\,{}^1K$ However, it is not enough to find a plausible theoretical explanation of only one doubly excited state in this neighbourhood. I consider it certain that the states ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ cannot be states with only one electron excited No satisfactory explanation on this basis is known of any of the ${}^1\Sigma_g$ states $3\,{}^1K$, ${}^1L$, ${}^1M$, ${}^1N$, $4\,{}^1K$, of the ${}^1\Pi_g$ state ${}^1Q$ or the upper state of the $\lambda 4143$ progression This list of probable doubly excited states may possibly be reduced ultimately by three I think it possible, though not likely, that ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ together are actually one genuine state + a collection of coincidences I regard it as certain that ${}^1\Sigma_g(a) + {}^1\Sigma_g(b)$ contain at least one real state, but beyond that one cannot go at present with complete confidence The states $4\,{}^1K$ and the upper states of the λ4143 progression might possibly be fragments of some singly excited states having unusual properties, though I think this very unlikely with regard to the last one I regard any further reductions than these as extremely unlikely

We are thus left with nine states which go down to $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ and which probably have both electrons excited In six of them the double excitation must be regarded as practically certain, in two others as highly probable and the last ($4\,{}^1K$) as rather doubtful All of these have their ν_e's (electronic levels) in the gap which extends from about 7000 to 23,000 cm $^{-1}$, roughly 1 to 3 V, below the ground level of the ion H_2^+ Even if we are not able to assign definite configurations to this collection of states with any certainty, it is legitimate to consider whether their existence is compatible with what is otherwise known about H_2 and what, if any, restrictions it would imply With this object I have compiled, in Table XI, a list of potential Rydberg series of doubly excited states whose earlier members will possibly, in some cases we can say will probably, lie rather deep, based on Teller's computations of the energies of the various excited states of the H_2^+ ion This table is based on the assumption that the coupling between the orbital and spin wave functions is negligible, this is very small for molecules with only one electron excited

Tablf XI

$2p\sigma\, 2p\sigma\, {}^1\Sigma_g$,	$2p\sigma\, 3p\sigma\, {}^1\Sigma_g$,	$2p\sigma\, 4p\sigma\, {}^1\Sigma_g$,	$2p\sigma\, {}^2\Sigma_u + e$
$2p\sigma\, 2p\pi\, {}^1\Pi_g$,	$2p\sigma\, 3p\pi\, {}^1\Pi_g$,	$2p\sigma\, 4p\pi\, {}^1\Pi_g$,	$2p\sigma\, {}^2\Sigma_u + e$
$2p\sigma\, 2p\pi\, {}^1\Pi_g$,	$2p\pi\, 3p\sigma\, {}^1\Pi_g$,	$2p\pi\, 4p\sigma\, {}^1\Pi_g$,	$2p\pi\, {}^2\Pi_u + e$
$2p\pi\, 2p\pi\, {}^1\Sigma_g$,	$2p\pi\, 3p\pi\, {}^1\Sigma_g$,	$2p\pi\, 4p\pi\, {}^1\Sigma_g$,	$2p\pi\, {}^2\Pi_u + e$
$2p\pi\, 2p\pi\, {}^1\Delta_g$,	$2p\pi\, 3p\pi\, {}^1\Delta_g$,	$2p\pi\, 4p\pi\, {}^1\Delta_g$,	$2p\pi\, {}^2\Pi_u + e$
$2p\sigma\, 2s\sigma\, {}^1\Sigma_u$,	$2p\sigma\, 3s\sigma\, {}^1\Sigma_u$,	$2p\sigma\, 4s\sigma\, {}^1\Sigma_u$,	$2p\sigma\, {}^2\Sigma_u + e$
$2s\sigma\, 2p\sigma\, {}^1\Sigma_u$,	$2s\sigma\, 3p\sigma\, {}^1\Sigma_u$,	$2s\sigma\, 4p\sigma\, {}^1\Sigma_u$	$2s\sigma\, {}^2\Sigma_g + e$
$2s\sigma\, 2s\sigma\, {}^1\Sigma_g$,	$2s\sigma\, 3s\sigma\, {}^1\Sigma_g$,	$2s\sigma\, 4s\sigma\, {}^1\Sigma_g$,	$2s\sigma\, {}^2\Sigma_g + e$
	$2p\sigma\, 3d\sigma\, {}^1\Sigma_u$,	$2p\sigma\, 4d\sigma\, {}^1\Sigma_u$,	$2p\sigma\, {}^2\Sigma_u + e$

The ionic states $2p\sigma\, {}^2\Sigma_u$ and $2s\sigma\, {}^2\Sigma_g$ are unstable, so that in series of the types which end up with these ions when $n = \infty$ the internuclear distance should tend to ∞ as n increases The ionic state $2p\pi\, {}^2\Pi_u$ is only just stable, having a very weak minimum in its energy *vs* internuclear distance curve at about $r_e = 4$ A This is more than twice as great as r_e for any known H_2 state, so that its effect is likely to be much the same as $r_e = \infty$ so far as the first two or three states of a given series is concerned

In attempting to fit the nine states into the array in Table XI we note first that none of them can be any of the odd, suffix u, states, since transitions from these to the odd $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ state are prohibited Two allocations seem to be immediately indicated. Among the nine experimental states there is only one $^1\Pi_g$ state and one $^1\Delta_g$ state. It seems obvious to assign ^{1}Q, the one $^1\Pi_g$ state, to $2p\sigma\, 2p\pi\, ^1\Pi_g$, the deepest $^1\Pi_g$ state, and the upper state of the $\lambda 4143$ progression, which can be either a $^1\Delta_g$ state, a $^1\Phi_g$ state or a $^1\Gamma_g$ state, to the deepest $^1\Delta_g$ state $2p\pi\, 2p\pi\, ^1\Delta_g$. This also seems to be right for another reason. On Weizel's argument the deepest state would be $2p\sigma\, 2p\sigma\, ^1\Sigma_g$ and would lie somewhere near the position of $^1\Sigma_g(a)$ or $^1\Sigma_g(b)$. If we identify $^1\Sigma_g(a) + {}^1\Sigma_g(b)$ with $2p\sigma\, 2p\sigma\, ^1\Sigma_g$ we expect the next state above this to be $2p\sigma\, 2p\pi\, ^1\Pi_g$, and actually ^{1}Q is the deepest of the states above $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$. This has brought us to our first difficulty, because $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ appear to be two distinct states, whereas the scheme only provides one $2p\sigma\, 2p\sigma\, ^1\Sigma_g$ state. This difficulty can be avoided by assuming, what is not impossible, that $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ are not two distinct states but a jumble of one real state with an assortment of coincidences. The properties and constants of $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ are all very much alike, and if they are a jumble the constants of the real state will be close to the average. On this basis they are effective quantum number, or denominator, $z \approx 2\cdot18_6$, vibration difference $\omega_0 - x\omega_0 \approx 2265$ cm $^{-1}$, internuclear distance $r_0 \approx 1\cdot55$ A. The next member of the series $2p\sigma\, 3p\sigma\, ^1\Sigma_g$ would be likely to have z about 1 higher, $\omega_0 - x\omega_0$ lower and r_0 higher. Both ^{1}M and ^{1}L satisfy these requirements, as may be seen from Table XII, where the more important relevant constants of the nine experimental states and the three theoretical ions which terminate the various series in Table XI are given, ω_0 and $x\omega_0$ are not given separately since, for several states, only the combination $\omega_0 - x\omega_0$ is known. Some of the values in the table can only be determined rather roughly, but they are probably good enough for this discussion. On this scheme it would be natural to choose ^{1}L for $2p\sigma\, 3p\sigma\, ^1\Sigma_g$ in preference to ^{1}M, as its constants are tending more rapidly towards the ion $2p\sigma\, ^2\Sigma_u$ which is the end of the series than those of ^{1}M, and as otherwise there is not much to choose between them. $3\,^1$K might quite well be $2p\pi\, 2p\pi\, ^1\Sigma_g$ and ^{1}N, $2p\pi\, 3p\pi\, ^1\Sigma_g$, these two are almost certain to be either the lowest states of two separate series or the two lowest states of a single series, as they give the strongest transitions down to $1s\sigma\, 2p\sigma\, ^1\Sigma_u$ of any of the nine experimental states. This leaves ^{1}M and $4\,^1$K, which might quite well be $2s\sigma\, 2s\sigma\, ^1\Sigma_g$ and $2s\sigma\, 3s\sigma\, ^1\Sigma_g$ (or $2s\sigma\, 3d\sigma\, ^1\Sigma_g$) respectively.

I admit that all this is rather guessing, but I think it is something to know to what extent the nine experimental states can be fitted into such a theoretical framework qualitatively without too great a probable strain on the unknown quantitative restrictions. The question cannot be settled definitely until the wave functions of these doubly excited states are worked out That is a task for which I am not well fitted and it appears that no one is much attracted by it What we can say is that all but one of the nine experimental states can be fitted qualitatively into this theoretical scheme without being quantitatively very improbable All we have to sacrifice is the position that each of the two states $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ is a distinct and separate state and, as we have seen, it is not certain that they are Even this sacrifice would not be necessary if one of the theoretical $^1\Sigma_g$ states should prove to lie unexpectedly deep, as there are nine members in the list of theoretical states which are expected to lie rather low as well as in the list of experimental states

There is one aspect of this problem that deserves further consideration If we take $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ at their face value as distinct states and consider the six $^1\Sigma_g$ states, $^1\Sigma_g(a)$, $^1\Sigma_g(b)$, 3 ^{1}K, ^{1}L, ^{1}M and ^{1}N, they all show a good deal of resemblance to one another and have characteristics which mark them off as a group from other $^1\Sigma$ states But there is something about them which is even more remarkable Out of the six states, four occur in two pairs the members of each of which are almost identical These are $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ on the one hand and ^{1}M and ^{1}L on the other $^1\Sigma_g(a)$ and $^1\Sigma_g(b)$ have practically the same electronic energies, their ν_e's only differing by about 227 cm $^{-1}$ The other constants are very much the same in each molecule The small, or at any rate not large, differences in the vibrational and rotational constants are such as would be expected for molecules which have an almost identical structure apart from a small but definite difference in the equilibrium internuclear distance Statements nearly the same as these can be made about the pair ^{1}M and ^{1}L The only outstanding difference between them and the pair $^1\Sigma_g(a)+{}^1\Sigma_g(b)$ is that, corresponding to the different ν_e's, their Rydberg denominators differ by 0 94 That they are states with a very close similarity is shown also by their values for the Mecke ratio B_0/ω_0, or rather $B_0/(\omega_0 - x\omega_0)$, which is the nearest we can get to it, in the absence of knowledge of the small fraction x for all of them More than forty singly excited states of H_2 are known and for each of them this ratio is not far from the average value 0 013 For the four states now being considered its range is 0 0055–0 0065 This suggests that the kind of binding is the same in all four states and much stronger than in the singly excited H_2 molecules All these facts point to the view that these four states are rather wide doublets forming two successive members of a Rydberg series This position is also supported by the fact that the doublet separation Δ_{ν_e} is greater, at about 227 cm $^{-1}$, for the lower member $^1\Sigma_g(a)+{}^1\Sigma_g(b)$ than it is, at about 135, for the higher member ^{1}L + ^{1}M

It is possible that such doublets could arise if there were strong inter-

TABLE XII

	${}^1\Sigma_g(a)$	${}^1\Sigma_g(b)$	$3\,{}^1K$	${}^1L$	${}^1M$	${}^1N$	$4\,{}^1$K	^{1}Q	λ4143 prog	$2p\sigma\,{}^3\Sigma_u$ or $2s\sigma\,{}^3\Sigma_u$	$2p\pi\,{}^3\Pi_u$
$z \rightarrow$	2 18_9	2 19_3	2 9115	3 1143	3 1331	3 403	3 91	2 866	?3 312	∞	—
$\omega_e - x\omega_0 \rightarrow$	2331 7	2198	2233	ca 1835	2176	1983 3	Unknown	ca 730	Unknown	0	Small
B_e	ca 15 1	ca 13 6	ca 11	ca 10	ca 13	ca 17 5	ca 19?	ca 16	ca 37	0	—
r_e	ca 1 5	ca 1 6	ca 1 74	ca 1 84	ca 1 6	ca 1 38	ca 1 33?	ca 1 45	ca 0 9_5	∞	ca 4
Λ	0	0	0	0	0	0	0	1	2, 3 or 4	0	0
$B_e/(\omega_e - x\omega_0)$	0 0065	0 0062	0 0049	0 0055	0 0060	0 0088	Unknown	0 022	Unknown	—	—
ν_e	23152	2292_5	12938	11308	11173	9468	ca 7160	13350	Probably ca 10000	—	—

coupling between orbital and spin angular momenta in these doubly excited states I do not profess to understand why this should be so but, after all, these states, and a few other states of H_2 about which we are not so well informed, are the only doubly excited molecular states known and there is no very definite theory about them, so that almost anything may be expected For example, in the theoretical state we have designated as $2p\pi\,2p\pi\,{}^1\Sigma_g$ the real quantum numbers, if separable, are 0, 0, ± 1, $\pm \frac{1}{2}$ The orbital angular momentum and spin can be combined to give the resultants either $\frac{3}{2}$ or $\frac{1}{2}$ This is a doublet with components ${}^1\Sigma_{\frac{1}{2}}$ and ${}^1\Sigma_{\frac{3}{2}}$ Why these states should lie so deep and why the doublets are so widely separated are both mysterious, but I think it would be at least as mysterious if the facts I have recited were all a coincidence On this view ${}^1Q$ could be $2p\pi\,2p\sigma\,{}^1\Pi_g$ and the upper state of the $\lambda 4143$ progression $2p\pi\,2p\pi\,{}^1\Delta_g$, each, as before ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ would be $2p\pi\,2p\pi\,{}^1\Sigma_{\frac{1}{2}}$ and $2p\pi\,2p\pi\,{}^1\Sigma_{\frac{3}{2}}$, whilst ${}^1L$ and ${}^1M$ would be $2p\pi\,3p\pi\,{}^1\Pi_{\frac{1}{2}}$ and $2p\pi\,3p\pi\,{}^1\Pi_{\frac{3}{2}}$ As a general rule the levels with $j = \frac{3}{2}$ should lie higher than those with $j = \frac{1}{2}$, but in this case I am inclined to associate ${}^1\Sigma_g(a)$ with ${}^1M$ and ${}^1\Sigma_g(b)$ with ${}^1L$ ${}^1\Sigma_g(a)$ and ${}^1M$ are the stronger components of the doublets, and the rotational and vibrational constants all run more like consecutive members of series with this arrangement It is likely that one of the doublets is inverted as the result of a perturbation We have seen that ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ are both perturbed The rotational structure of ${}^1M$ (Richardson and Davidson 1929, p 62) looks as if this state were perturbed, and unless Richardson and Davidson's (1929, p 57) interpretation of ${}^1L$ is altogether wrong ${}^1L$ must be strongly perturbed as the $v' = 1$ level of ${}^1L \rightarrow 1s\sigma\,2p\sigma\,{}^1\Sigma_u$ has never yet been found

This arrangement makes $2p\pi\,2p\pi\,{}^1\Sigma_{\frac{1}{2},\frac{3}{2}}$ the deepest of the doubly excited states of H_2 and ${}^1Q$, which is the next deepest state, has been identified with $2p\sigma\,2p\pi\,{}^1\Pi_g$ This makes it reasonable to expect $3\,{}^1K$, the next deepest state to these, to be $2p\sigma\,2p\sigma\,{}^1\Sigma_g$, an identification I proposed in 1932 (Richardson 1934) This leaves us with only ${}^1N$ and $4\,{}^1K$ ${}^1N$ may very well prove to be $2s\sigma\,2s\sigma\,{}^1\Sigma_g$ and $4\,{}^1K$, $2p\sigma\,3p\sigma\,{}^1\Sigma_g$ Each of these has a closer affinity to a normal Σ state than the others and, on this view, this might be expected

I have now described two possible alternative configurations for the nine abnormal states which go down to $1s\sigma\,2p\sigma\,{}^1\Sigma_u$ The first of these is probably more in accordance with current theoretical ideas, but I think the second is a better description of the experimental facts, certainly if both the new states ${}^1\Sigma_g(a)$ and ${}^1\Sigma_g(b)$ are genuine

Except in H_2 no states with more than one electron excited have been recognized, but I know of no reason why they should not exist in other molecules I should, in fact, not be surprised if they turned out ultimately to be of not uncommon occurrence

SUMMARY

This paper describes two new band systems which go down to $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$, the state on which the strongest hitherto known systems in the visible end The systems extend from the yellow to the photographic infra-red limit (11750 A) and probably beyond The upper states are both ${}^1\Sigma_g$ states They are very close to one another and appear to be perturbing each other as well as the state $1s\sigma\, 2s\sigma\, {}^1\Sigma_g$, which is in the same neighbourhood Their rotational structure resembles closely that of the "abnormal" ${}^1\Sigma_g$ states $3\,{}^1K$, ${}^1L$, ${}^1M$ and ${}^1N$ Reasons are given which make it impossible to regard either of the two new states as states with only one electron excited Two alternative methods of classifying the two new states and the seven other abnormal states which have transitions to $1s\sigma\, 2p\sigma\, {}^1\Sigma_u$ and uniting them into a consistent scheme as states with both electrons excited are put forward and their merits discussed

REFERENCES

Allibone 1926 *Proc Roy Soc* A, **112**, 196–213
Davidson 1932 *Proc Roy Soc* A, **138**, 580–93
Deodhar 1926 *Proc Roy Soc* A, **113**, 420–32
Dieke 1929 *Z Phys* **55**, 447–50
— 1936*a* *Phys Rev* **50**, 797–805
— 1936*b* Private communication
Gale, Monk and Lee 1928 *Astrophys J* **67**, 89–113
Kapuscinski and Eymers 1929 *Proc Roy Soc* A, **122**, 58–68
Merton and Barratt 1922 *Philos Trans* A, **222**, 369–400
Poetker 1927 *Phys Rev* **30**, 418–28
Richardson 1934 "Molecular Hydrogen and its Spectrum" New Haven Yale University Press
— 1937 *Proc Roy Soc* A, **160**, 487–507
Richardson and Davidson 1929 *Proc Roy Soc* A, **124**, 50–89
Richardson and Rymer 1934*a* *Proc Roy Soc* A, **147**, 251–72
— — 1934*b* *Proc Roy Soc* A, **147**, 272–92
Tanaka 1925 *Proc Roy Soc* A, **108**, 592–606
Teller 1930 *Z Phys* **61**, 458–80
Weizel 1930 *Z Phys* **65**, 456–63

The two-dimensional hydrodynamical theory of moving aerofoils—II

By Rosa M Morris, B Sc

Research Student, University College, Cardiff

(*Communicated by L N G Filon, F R S —Received 2 November* 1937)

1—Introduction

1 1—In the first paper of this series (referred to subsequently as "I") I developed in some detail the complete solution of the hydrodynamical problem of the motion of an incompressible homogeneous inviscid liquid when a cylinder with a general aerofoil cross-section is moving in any manner perpendicular to its axis As was mentioned in that paper, the discussion there suffers from two well-known defects In the first place it fails generally when there is a sharp edge to the moving cylinder—or a singularity on the boundary curve—because the velocity of the liquid, as defined by the potential function, becomes infinite at such an edge This difficulty is usually surmounted by a proper choice of the circulation, but in the case of the moving cylinder this artifice is unavailing by itself as it only gives one condition, when the complete vanishing of the velocity at the trailing edge involves two In any case, however, the existence of finite circulation creates difficulties of its own—the energy involved in such circulations is infinite—so that still further complication of the mathematical problem seems required In practice it is known that the liquid motion is more complicated than that represented in the simpler problem, as it is always accompanied by the development, at such a sharp edge, and particularly in unsteady motion, of a region of turbulent motion in the fluid which trails behind in the wake of the cylinder To deal with this wake and its effects mathematically we must, of course, make certain simplifying assumptions The simplest picture is that used by Wagner (1925), who imagined the wake to be a simple surface of discontinuity (vortex sheet) trailing behind in the liquid Of course, in a perfect fluid such a sheet could not arise, but nevertheless, by assuming its existence, it is possible without any close inquiry into its structure to calculate its effect on the forces acting on the cylinder This is, in effect, what Wagner did, but under conditions which imply that his results are only approximately true when the cylinder is a flat plate with very small motion (without rotation)

The main objects of the present paper are, first, to show that the general theory of the first paper can be extended to include the effects of such surfaces of discontinuity and, second, to discuss the bearing of the general results obtained in the particular problems discussed by Wagner, and others in extension of his work, and thereby to test the validity of the formulae now used in practice, the various terms in which have been obtained by diverse devices, all indirect and often involving contradictory physical assumptions and approximations

1 2—To obtain a complete comparison with the known theoretical results we have included here also an analysis, on the lines of the previous paper, of the more usual and familiar statical problem, that is, the case where the cylinder is at rest in a uniform stream In order to facilitate the discussion of the trailing wake problem, we have also solved as a preliminary step the simple case of the motion outside the cylinder at rest, which arises from the presence in the fluid of a single-line vortex parallel to its generators It may not be useless to emphasize the directness of the solutions obtainable by the new method of approach Whereas, in the past, practically all results in this theory have been obtained by indirect methods involving at times a confusing mixture of mathematics and physics, the new method makes a direct division between the two stages of any problem of this character, the first, the analytical or geometrical problem of the transformation of the cylinder cross-section on to a circle, which in the present discussion is assumed accomplished, and the second, the purely physical problem when this cylinder is the boundary

Full references to the existing work in this subject are given in detail in the second volume of *Aerodynamic Theory* by Durand (1935), and our own references are mainly to the various sections in that book which deal with the specific problems we have under review

2—THE GENERAL AEROFOIL IN A UNIFORM STREAM

2 1—The aerofoil is considered as existing in a uniform stream in which the undisturbed velocity components, referred to any rectangular axes, are U, V In addition we shall consider the effects of a circulation of intensity K round the cylinder

Using the notation of "I", the undisturbed potential is

$$\Omega_0 = \overline{W}z,$$

where $$W = U + iV \quad \text{and} \quad \overline{W} = U - iV$$

The general transformation for the aerofoil being

$$z = c \sum_{n=0}^{\infty} a_n e^{(n-1)i\zeta},$$

we have Ω_0 in terms of ζ as follows

$$\Omega_0 = \overline{W} c \sum_{n=0}^{\infty} a_n e^{(n-1)i\zeta}$$

When the cylinder is introduced into the stream the new potential function must be such that it reduces to its original form at infinity, $\eta = +\infty$, and must also be purely real on the surface of the cylinder, $\eta = 0$, which is the stream line $\psi = 0$

Now Ω_0 contains two types of terms

$$F_1(\zeta) = \overline{W} c \sum_{n=1}^{\infty} {}_n e^{(n-1)i\zeta}$$

and $$F_2(\zeta) = \overline{W} c a_0 e^{-i\zeta},$$

which are distinguished by their behaviour at infinity $F_1(\zeta)$ is such that it tends to zero at infinity and has a conjugate which diverges there, whilst $F_2(\zeta)$ diverges at infinity and has a conjugate which tends to zero there To form a new potential function Ω from Ω_0, which is purely real when $\eta = 0$, we might simply add the conjugate of all the terms in Ω_0, but the additional terms must reduce to zero at infinity, where the presence of the cylinder is unnoticed If we add, however, the conjugate of $F_1(\zeta)$ we are adding terms which diverge at infinity, so we must adopt the alternative and remove $F_1(\zeta)$ altogether, and add the conjugate of $F_2(\zeta)$ We write therefore

$$\begin{aligned} \Omega &= \Omega_0 - F_1(\zeta) + \overline{F}_2(\zeta) \\ &= \overline{W} c a_0 e^{-i\zeta} + W c \bar{a}_0 e^{i\zeta}, \end{aligned}$$

which is a potential function satisfying all the conditions of the problem It may be of some interest to notice that it only contains the first two coefficients in the expansion of z, so that in effect it is practically the same for all types of cylinder

2 2—With a circulation K round the aerofoil we then have

$$\Omega = \overline{W} c a_0 e^{-i\zeta} + W c \bar{a}_0 e^{i\zeta} - \frac{K\zeta}{2\pi},$$

and now we can calculate the components X, Y of the force on the aerofoil by means of the usual contour integral formulae

$$Y+\imath X=-\tfrac{1}{2}\rho\int_{c}\left(\frac{d\Omega}{dz}\right)^{2}dz=\tfrac{1}{2}\rho\int_{0}^{2\pi}\left(\frac{d\Omega}{d\zeta}\right)^{2}\frac{d\zeta}{dz/d\zeta},$$

where $$\frac{d\Omega}{d\zeta}=-\imath\overline{W}ca_{0}e^{-\imath\zeta}+\imath Wc\bar{a}_{0}e^{\imath\zeta}-\frac{K}{2\pi}$$

and $$\left(\frac{dz}{d\zeta}\right)^{-1}=\frac{\imath e^{\imath\zeta}}{ca_{0}}\left\{1-\frac{a_{2}}{a_{0}}e^{2\imath\zeta}-\frac{2a_{3}}{a_{0}}e^{3\imath\zeta}-\quad\right\}^{-1}$$

The change of sign in converting the integral from the z to the ζ-plane arises from the fact that the convention we have adopted that $\eta=+\infty$ corresponds to $z=+\infty$ makes the direction of the circulation in the two planes opposite to one another This accounts also for the unusual sign in the term in K *

Picking out the constant term in the product of $(d\Omega/d\zeta)^{2}$ and $(dz/d\zeta)^{-1}$, which is the only term which contributes anything in the integral, we derive for the complex force

$$Y+\imath X=-\rho K\overline{W},$$

the familiar Kutta-Joukowski formula

The couple is similarly the real part of

$$-\tfrac{1}{2}\rho\int_{c}\left(\frac{d\Omega}{dz}\right)^{2}z\,dz=\tfrac{1}{2}\rho\int_{0}^{2\pi}\left(\frac{d\Omega}{d\zeta}\right)^{2}\frac{z\,d\zeta}{dz/d\zeta},$$

and in the same way this proves to be the real part of

$$\frac{\rho\imath K^{2}}{4\pi}-\rho K\overline{W}ca_{1}-2\pi\rho\imath c^{2}\overline{W}^{2}a_{0}a_{2}+2\pi\rho\imath c^{2}W\overline{W}a_{0}a_{0}$$

The first and last terms are of course purely imaginary and contribute nothing to the couple

3—THE LINE VORTEX OUTSIDE THE CYLINDER

3 1—Let us now consider the motion in the liquid surrounding the cylinder when a line vortex is situated outside the cylinder, say at the point z_s, thus generalizing some results obtained by Bickley (1929) for the simple case of a circular cylinder

---

* The necessity for this change of sign was pointed out to me by Professor Filon it also affects the work of "I", where, in fact, the results all appear with the wrong sign

The undisturbed velocity potential of the stream due to the line vortex of strength m passing through the point z_s is

$$\Omega_0 = -(im/2\pi)\log(z-z_s)$$

The cylinder being defined as before by the transformation

$$z = c\sum_{n=0}^{\infty} a_n e^{(n-1)i\zeta}, \quad \eta = 0,$$

the undisturbed potential in terms of ζ will be

$$\Omega_0 = -(im/2\pi)\log c\{(a_0 e^{-i\zeta}+a_1+a_2 e^{i\zeta}+\quad)-(a_0 e^{-i\zeta_s}+a_1+a_2 e^{i\zeta_s}+\quad)\},$$

where $\zeta_s = \xi_s + i\eta_s$ is the point in the ζ-plane corresponding to z_s This can be written in the form

$$\Omega_0 = \text{const} \; -(im/2\pi)\log\{1-e^{-i(\zeta-\zeta_s)}\}-(im/2\pi)\log F(\zeta),$$

where $F(\zeta)$ is a polynomial in $\exp(i\zeta)$ whose conjugate diverges at infinity Proceeding along the lines sketched above we find the new potential function when the cylinder is placed in the stream to be simply

$$\Omega = \text{const} \; -(im/2\pi)\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\}$$

The force and couple on the cylinder depend on the value of $d\Omega/d\zeta$, which in this case is found to be

$$-\frac{m}{2\pi}\left[1-\frac{\sinh\eta_s}{\cosh\eta_s-\cos(\zeta-\xi_s)}\right]$$

3 2—The components X, Y of the force on the cylinder are again given by

$$Y+iX = -\tfrac{1}{2}\rho\int_c\left(\frac{d\Omega}{dz}\right)^2 dz = \tfrac{1}{2}\rho\int_0^{2\pi}\left(\frac{d\Omega}{d\zeta}\right)^2\frac{d\zeta}{dz/d\zeta},$$

so that we want the value of

$$\int_0^{2\pi}\left[1-\frac{\sinh\eta_s}{\cosh\eta_s-\cos(\zeta-\xi_s)}\right]^2\frac{d\zeta}{dz/d\zeta}$$

This integral can be completed in a simple contour integral in the ζ-plane, round the rectangular contour with sides $\xi = 0$, $\eta = 0$, $\xi = 2\pi$, $\eta = \infty$ the integrand has the same value at two corresponding points $i\eta$ and $2\pi + i\eta$, so that the integrals along $\xi = 0$ and $\xi = 2\pi$ cancel out The behaviour at infinity is also obviously regular, so that the value of the integral is obtained directly from the residues at the poles contained in the contour But the only pole inside is at $\zeta = \zeta_s$, and thus we have to evaluate the residue there Writing $\zeta = \zeta_s + t$, where t is small, we have approximately

$$\left[1 - \frac{\sinh \eta_s}{\cosh \eta_s - \cos(\zeta - \xi_s)}\right]^2 = -\frac{1}{t^2}[1 - 2it(1 - \tfrac{1}{2}\coth \eta_s)], \tag{i}$$

whilst since $$(dz/d\zeta) = z_s' + z_s'' t,$$

we have $$(dz/d\zeta)^{-1} = \frac{1}{z_s'}\left(1 - \frac{z_s''}{z_s'} t\right), \tag{ii}$$

where dashes denote differentiation with respect to ζ, so that

$$z' = ic\sum_{n=0}^{\infty}(n-1)\,a_n e^{(n-1)i\zeta},$$

$$z'' = -c\sum_{n=0}^{\infty}(n-1)^2 a_n e^{(n-1)i\zeta}$$

To evaluate the residue of the integral we want simply the coefficient of t^{-1} in the product of the two expansions (i) and (ii) and this is

$$\frac{1}{z_s'}\left\{\frac{z_s''}{z_s'} + 2i - i\coth \eta_s\right\}$$

The components of the force are thus given by

$$Y + iX = \frac{m^2\rho}{4\pi z_s'}\left\{\frac{iz_s''}{z_s'} - 2 + \coth \eta_s\right\}$$

This result can be recast in a form which displays its full physical significance

The actual velocity of the liquid in the neighbourhood of the vortex, excluding the large motions due to the swirling of the vortex itself, is

$$\overline{W}_s = \frac{m}{2\pi z_s'}\left\{\frac{e^{i(\zeta_s - \bar{\zeta}_s)}}{1 - e^{i(\zeta_s - \bar{\zeta}_s)}}\right\} = \frac{m}{4\pi z_s'}\{\coth \eta_s - 1\}$$

Thus the components of the force have the value

$$\rho m \overline{W}_s - \frac{m^2\rho}{4\pi z_s'}\left(1 - \frac{iz_s''}{z_s'}\right)$$

The first or main part of the force represents the reaction between the stream of the vortex and the stream induced by the cylinder, whilst the second is the purely intrinsic term which is actually equal to the centrifugal force on the mass of liquid which occupies the space of the cylinder in the free vortex stream

3 3—In reality the vortex will move with the liquid in its neighbourhood unless constrained somehow to remain at rest But, if the vortex moves, there are other terms in the pressure equation, and these contribute to the expression for the components of the force a term

$$-\tfrac{1}{2}\rho\int_c\left(\frac{\partial\Omega}{\partial t}+\frac{\overline{\partial\Omega}}{\partial t}\right)dz$$

Now $$\Omega=-\frac{im}{2\pi}\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\zeta_s)}}\right\},$$

so that $$\frac{\partial\Omega}{\partial t}+\frac{\overline{\partial\Omega}}{\partial t}=-\frac{m}{\pi}\left\{\zeta_s\frac{e^{-i(\zeta-\zeta_s)}}{1-e^{-i(\zeta-\zeta_s)}}+\dot{\zeta}_s\frac{e^{i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\zeta_s)}}\right\}$$

On the cylinder the exponentials are such that their moduli are less than unity, and we can expand the denominators, giving

$$\frac{\partial\Omega}{\partial t}+\frac{\overline{\partial\Omega}}{\partial t}=-\frac{m}{\pi}\left\{\zeta_s\sum_{n=1}^{\infty}e^{-ni(\zeta-\zeta_s)}+\dot{\zeta}_s\sum_{n=1}^{\infty}e^{ni(\zeta-\zeta_s)}\right\}$$

We also have $$d\bar{z}=\left\{-ic\sum_{n=0}^{\infty}(n-1)\,\bar{a}_n e^{-(n-1)i\zeta}\right\}d\zeta,$$

so that the integration of the product round the contour of the cylinder—given by the constant term in the integrand alone—is simply

$$\begin{aligned}&-m\rho\left[\zeta_s ic\bar{a}_0e^{i\zeta_s}+\dot{\zeta}_s\left\{-ic\sum_{n=2}^{\infty}(n-1)\,\bar{a}_n e^{-(n-1)i\zeta_s}\right\}\right]\\&=-m\rho[\zeta_s ic\bar{a}_0e^{i\zeta_s}-\dot{\zeta}_s\{ic\bar{a}_0e^{i\bar{\zeta}_s}-\bar{z}'_s\}]\\&=-m\rho[\zeta_s ic\bar{a}_0e^{i\zeta_s}-\dot{\zeta}_s\{ic\bar{a}_0e^{i\bar{\zeta}_s}+\overline{W}_s\}]\end{aligned}$$

Thus in all $$Y+iX=-m\rho ic\bar{a}_0[\zeta_s e^{i\zeta_s}-\dot{\zeta}_s e^{i\zeta_s}]-\frac{m^2\rho}{4\pi z'_s}\left[1-\frac{iz''_s}{z'_s}\right]$$

3 4—To find the steady motion terms in the couple on the cylinder we again use the contour integral The couple is as we have seen, the real part of

$$-\tfrac{1}{2}\rho\int_c\left(\frac{\partial\Omega}{\partial z}\right)^2 z\,dz=\tfrac{1}{2}\rho\int_0^{2\pi}\left(\frac{\partial\Omega}{\partial\zeta}\right)^2\frac{z\,d\zeta}{dz/d\zeta}$$

The integral being completed in the usual contour integral in the ζ-plane, the only pole is at $\zeta=\zeta_s$, and we have again to evaluate the residue there. Writing $\zeta = \zeta_s + t$, where t is small, the above integral is approximately

$$-\frac{\rho m^2}{8\pi^2}\int \frac{1}{t^2}[1 - 2it(1 - \tfrac{1}{2}\coth \eta_s)]\frac{(z_s + z_s' t)}{(z_s' + z_s'' t)}\,dt$$

$$= -\frac{\rho m^2}{8\pi^2}\frac{z_s}{z_s'}\int \frac{1}{t^2}[1 - 2it(1 - \tfrac{1}{2}\coth \eta_s)]\left\{1 + \frac{z_s' t}{z_s} - \frac{z_s'' t}{z_s'}\right\}dt$$

Apart from the factor z_s which is present we see that the only additional term introduced is imaginary and therefore does not affect the value of the couple Our result therefore is that the couple is the real part of

$$z_s(Y + iX),$$

just as if the force is applied from the position of the vortex In the above case the steady motion terms in the couple are therefore the real part of

$$\rho m z_s \overline{W}_s - \frac{m^2 \rho z_s}{4\pi z_s'}\left(1 - \frac{i z_s''}{z_s'}\right)$$

The acceleration terms due to the velocity of the vortex are the real part of

$$-\tfrac{1}{2}\rho\int_c \left(\frac{\partial \Omega}{\partial t} + \overline{\frac{\partial \Omega}{\partial t}}\right) z\,dz,$$

where $$\frac{\partial \Omega}{\partial t} + \overline{\frac{\partial \Omega}{\partial t}} = -\frac{m}{\pi}\left\{\zeta_s \sum_{n=1}^{\infty} e^{ni(\zeta - \zeta_s)} + \bar{\zeta}_s \sum_{n=1}^{\infty} e^{-ni(\zeta - \bar{\zeta}_s)}\right\}$$

The real part of the above will therefore be

$$-\frac{m\rho}{4\pi}\int_0^{2\pi}\left\{\zeta_s \sum_{n=1}^{\infty} e^{ni(\zeta - \zeta_s)} + \bar{\zeta}_s \sum_{n=1}^{\infty} e^{-ni(\zeta - \bar{\zeta}_s)}\right\}\frac{d(zz)}{d\zeta}\,d\zeta$$

But with the notation of "I" we have, on the cylinder,

$$\frac{d(z\bar{z})}{d\zeta} = i\left\{\sum_{n=1}^{\infty} n b_n e^{ni\zeta} - \sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}}\right\},$$

where $$b_n = \sum_{r=0}^{\infty} a_{r+n}\bar{a}_r,$$

so that the acceleration terms are simply

$$-\tfrac{1}{2}m\rho i\left\{\zeta_s \sum_{n=1}^{\infty} n b_n e^{ni\zeta_s} - \bar{\zeta}_s \sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}_s}\right\}$$

Taking only the real part, the couple is therefore

$$\tfrac{1}{2}\rho m\{z_s\overline{W}_s+\bar{z}_s W_s\}-\frac{m^2\rho z_s}{8\pi z_s'}\left\{1-\frac{iz_s''}{z_s'}\right\}$$
$$-\frac{m^2\rho\bar{z}_s}{8\pi\bar{z}_s'}\left\{1+\frac{i\bar{z}_s''}{\bar{z}_s'}\right\}-\tfrac{1}{2}m\rho i\left\{\zeta_s\sum_{n=1}^{\infty}nb_n e^{ni\zeta_s}-\bar{\zeta}_s\sum_{n=1}^{\infty}nb_n e^{-ni\bar{\zeta}_s}\right\}$$

It is easy to see that these results agree with those obtained by Bickley for the circular cylinder

4—THE TRAILING WAKE PROBLEM

4 1—In this paragraph we shall give a general formulation of Wagner's discussion of the effect of a surface of discontinuity (vortex sheet) on the motion of a general cylinder

We assume then that the cylinder is moving with velocity components u, v and angular velocity ω, that there is a circulation K round it, and a trail of vortices forming a vortex sheet with a section along a curve s in the co-ordinate plane, from a point $z = z_0$ to another point $z = z_1$ We shall assume that the strength of this sheet is m_s at the typical point, so that the element ds of the curve about that point is equivalent, in effect, to a line vortex of strength $m_s ds$ Owing to the difficulty arising from the existence of a finite circulation extending to infinity, we shall assume that the total circulation associated with the sheet just balances the circulation round the cylinder, or that

$$K+\int_{z_0}^{z_1} m_s ds = 0$$

The complete velocity potential of the fluid motion can of course be obtained by superposing the velocity potential of the vortex motion on that given in "I" for the remaining liquid motion it is therefore

$$\Omega = \Omega_a-\frac{K\zeta}{2\pi}-\int_{z_0}^{z_1}\frac{im_s}{2\pi}\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\}ds,$$

where Ω_a is the velocity potential, obtained in "I", of the fluid motion due to the cylinder alone, without circulation or vorticity and ζ_s denotes the value of ζ at the typical point on the wake curve

We must now choose the circulations in such a way that there is zero or finite velocity of the liquid at the trailing edge mathematically this means that $d\Omega/d\zeta$ must be zero at the same point of the boundary curve as $dz/d\zeta$ If this point is $\zeta = \zeta_t$, this means that

$$\left(\frac{d\Omega_a}{d\zeta}\right)_{\zeta=\zeta_t} - \frac{K}{2\pi} - \int_{s_0}^{s_1} \frac{m_s}{2\pi}\left\{1 - \frac{\sinh\eta_s}{\cosh\eta_s - \cos(\zeta_t - \xi_s)}\right\} ds = 0,$$

or, using the condition of zero total circulation,

$$\left(\frac{d\Omega_a}{d\zeta}\right)_{\zeta=\zeta_t} = -\frac{1}{2\pi}\int_{s_0}^{s_1}\left\{\frac{m_s \sinh\eta_s}{\cosh\eta_s - \cos(\zeta_t - \xi_s)}\right\} ds$$

This equation defines the dependence of the vorticity in the discontinuity surface on the motion of the cylinder it must, however, be emphasized that in reality, like every complex equation, it involves a double condition This point is important, as it shows that it is impossible to satisfy the condition of finite velocity at a sharp trailing edge in general by assuming a simple circulation

4 2—To calculate the components X, Y of the force on the cylinder we use our previous formula for the pressure

$$p = p_0 - \tfrac{1}{2}\rho\left\{\frac{\partial\Omega}{\partial t} + \frac{\partial\overline{\Omega}}{\partial t} + \frac{\partial\Omega}{\partial z}\frac{\partial\overline{\Omega}}{\partial \bar{z}}\right\},$$

and then

$$Y + iX = \int_c p\, d\bar{z},$$

and the couple about the origin, in the same way, is the real part of

$$\int_c p z\, d\bar{z}$$

As we have seen in "I", when the conditions are steady, the contribution of the fluid pressures to the components of the force is given by

$$Y + iX = \tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z})\, d\bar{z} - \tfrac{1}{2}\rho\int_c\left(\bar{w} - i\omega\bar{z} - \frac{d\Omega}{dz}\right)^2 dz,$$

and the couple by the real part of

$$\tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z})\, z\, d\bar{z} - \tfrac{1}{2}\rho\int_c\left(\bar{w} - i\omega\bar{z} - \frac{d\Omega}{dz}\right)^2 z\, dz$$

But these two forms have to be generalized when there are accelerations by the addition of terms in the accelerations and variations in the circulation, and the strengths and positions of the vortices The acceleration terms and that due to a rate of change of circulation are identical with those obtained in "I", so we have merely to evaluate now the forces due to the variation in the strengths and positions of the vortices, these all being superposable

4 3—Consider first a single vortex of strength $m_s ds$ at the point ζ_s The extra acceleration terms in the integral for the components of the force will be

$$-\tfrac{1}{2}\rho\int_c\left\{\frac{\partial\Omega}{\partial\zeta_s}\zeta_s+\overline{\frac{\partial\Omega}{\partial\zeta_s}}\dot{\bar{\zeta}}_s+\frac{\partial\Omega}{\partial\bar{\zeta}_s}\dot{\bar{\zeta}}_s+\overline{\frac{\partial\Omega}{\partial\bar{\zeta}_s}}\zeta_s+\frac{\partial\Omega}{\partial m_s}m_s+\overline{\frac{\partial\Omega}{\partial m_s}}m_s\right\}dz$$

In our case this becomes, as before,

$$\frac{\rho m_s ds}{2\pi}\int_c\left\{\zeta_s\frac{e^{-\imath(\zeta-\zeta_s)}}{1-e^{-\imath(\zeta-\zeta_s)}}+\dot{\bar{\zeta}}_s\frac{e^{\imath(\zeta-\bar{\zeta}_s)}}{1-e^{-\imath(\zeta-\bar{\zeta}_s)}}\right\}d\bar{z}$$

$$+\frac{\rho\imath m_s ds}{2\pi}\int_c\log\left\{\frac{1-e^{-\imath(\zeta-\zeta_s)}}{1-e^{\imath(\zeta-\bar{\zeta}_s)}}\right\}d\bar{z}$$

We have seen in paragraph 3 3 that the first integral gives

$$-\rho m_s ds[\zeta_s\imath c\bar{a}_0e^{\imath\zeta_s}-\dot{\bar{\zeta}}_s\{\imath c\bar{a}_0e^{\imath\bar{\zeta}_s}+\dot{\bar{z}}_s\}]$$

In the second integral the exponentials are again such that their moduli are less than unity, on the cylinder, and we can therefore expand the logarithm in the form

$$\sum_{n=1}^{\infty}\frac{1}{n}\{e^{n\imath(\zeta-\bar{\zeta}_s)}-e^{-n\imath(\zeta-\zeta_s)}\}$$

Also, on the cylinder,

$$\frac{d\bar{z}}{d\zeta}=-\imath c\sum_{n=0}^{\infty}(n-1)\,a_n e^{-(n-1)\imath\zeta},$$

so that the constant term in the product is

$$-\imath c\left\{\bar{a}_0e^{\imath\zeta_s}+\sum_{n=2}^{\infty}\bar{a}_n e^{-(n-1)i\zeta_s}\right\}$$

or $$-\imath\{z_s-c\bar{a}_1-c\bar{a}_0e^{\imath\bar{\zeta}_s}+c\bar{a}_0e^{\imath\zeta_s}\}$$

The integral is therefore

$$-\rho m_s ds\{\bar{z}_s-c\bar{a}_1-c\bar{a}_0e^{\imath\bar{\zeta}_s}+c\bar{a}_0e^{\imath\zeta_s}\}$$

Thus, assuming a trail of vortices, the contribution to the components of the force due to the variations in the positions and strengths of the vortices will be

$$-\rho\int_{s_0}^{s_1}m_s[\dot{\bar{z}}_s+\imath c\bar{a}_0(\dot{\zeta}_s e^{\imath\zeta_s}-\dot{\bar{\zeta}}_s e^{\imath\bar{\zeta}_s})]\,ds,$$

$$-\rho\int_{s_0}^{s_1}m_s[\bar{z}_s-c\bar{a}_1-c\bar{a}_0e^{\imath\bar{\zeta}_s}+c\bar{a}_0e^{\imath\zeta_s}]\,ds$$

4 4—We have now to calculate the steady motion terms These are, as we have seen, given by

$$\tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z})\,d\bar{z} - \tfrac{1}{2}\rho\int_c \left(\bar{w} - i\omega\bar{z} - \frac{d\Omega}{dz}\right)^2 dz,$$

where $$\frac{d\Omega}{d\zeta} = \frac{d\Omega_a}{d\zeta} - \frac{K}{2\pi} - \int_{s_0}^{s_1} \frac{m_s}{2\pi}\left\{1 - \frac{\sinh\eta_s}{\cosh\eta_s - \cos(\zeta - \xi_s)}\right\} ds$$

The terms $$\tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z})\,d\bar{z}$$

and $$-\tfrac{1}{2}\rho\int_c \left[(\bar{w} - i\omega z)^2 - 2(\bar{w} - i\omega\bar{z})\frac{d\Omega_a}{dz}\right] dz$$

are calculated as in "I" and give the same results as there obtained, except that, as noticed before, the signs must be reversed The terms introduced by the circulation and wake are of two types There are first the two terms

$$\rho\int_0^{2\pi} (\bar{w} - i\omega\dot{z})\frac{K}{2\pi}\,d\zeta$$

and $$-\rho\int_{z_0}^{z_1} \frac{m_s\,ds}{2\pi}\int_0^{2\pi}\left[\frac{e^{-i(\zeta-\zeta_s)}}{1 - e^{-i(\zeta-\zeta_s)}} + \frac{e^{i(\zeta-\zeta_s)}}{1 - e^{i(\zeta-\zeta_s)}}\right](\bar{w} - i\omega z)\,d\zeta$$

The first integral is easily evaluated, whilst in the second the two exponentials both have, on the cylinder, where ζ is real, a modulus less than unity, so that the quantity in square brackets can be expanded in series Using also the form

$$\bar{z} = c\{\bar{a}_0 e^{i\zeta} + \bar{a}_1 + \bar{a}_2 e^{-i\zeta} + \quad\},$$

we see at once that the integral with respect to ζ in the second term is simply

$$-2\pi i\omega c(\bar{a}_0 e^{i\zeta_s} + \bar{a}_1 + \bar{a}_2 e^{-i\zeta_s} + \quad) = -2\pi i\omega(\bar{z}_s - c\bar{a}_0 e^{i\bar{\zeta}_s} - c\bar{a}_1 + c\bar{a}_0 e^{i\zeta_s}),$$

so that altogether this term contributes to the complex force

$$\rho\bar{w}K - \rho i c\omega K\bar{a}_1 + \rho i\omega\int_{z_0}^{z_1} m_s(\bar{z}_s - ca_0 e^{i\bar{\zeta}_s} - c\bar{a}_1 + c\bar{a}_0 e^{i\zeta_s})\,ds$$

Finally, there is the term

$$-\tfrac{1}{2}\rho\int_c \left(\frac{d\Omega}{dz}\right)^2 dz = \tfrac{1}{2}\rho\int_0^{2\pi}\left(\frac{d\Omega}{d\zeta}\right)^2 \frac{d\zeta}{dz/d\zeta},$$

keeping now all the terms in Ω together for convenience This integral in "I" was assumed to vanish Now however $d\Omega/dz$ has poles all along the wake, and at the zero of $dz/d\zeta$ We have however so chosen the wake strength and the circulation that $d\Omega/d\zeta$ is zero when $dz/d\zeta$ is zero, so that the latter pole has been removed To calculate the value of the integral we treat it again as a contour integral in the ζ-plane—the contour being a rectangle with sides $\xi = 0$, $\eta = 0$, $\xi = 2\pi$, $\eta = \infty$—and then the value is equal to the sum of the residues at the contained poles, that is, along the trace in this plane of the wake curve To find the sum of the residues our simplest plan is to surround the wake-curve trace by a double line along it, up on one side and down on the other, and in this way it is seen that the value of the integral is

$$\tfrac{1}{2}\rho\int_{z_0}^{z_1}\left\{\left(\frac{d\Omega_1}{dz}\right)^2-\left(\frac{d\Omega_2}{dz}\right)^2\right\}\frac{dz}{ds}\,ds,$$

Ω_1 being the value on one side of the wake and Ω_2 on the other side But on the wake

$$\frac{d\Omega_1}{ds}-\frac{d\Omega_2}{ds}=m_s, \quad \frac{d\Omega_1}{dn}=\frac{d\Omega_2}{dn},$$

so that

$$\frac{d\Omega_1}{dz}-\frac{d\Omega_2}{dz}=m_s\frac{ds}{dz}$$

The above integral is therefore simply

$$\tfrac{1}{2}\rho\int_{z_0}^{z_1}m_s\left(\frac{d\Omega_1}{dz}+\frac{d\Omega_2}{dz}\right)ds$$

But $\frac{1}{2}\left(\frac{d\Omega_1}{dz}+\frac{d\Omega_2}{dz}\right)$ is the conjugate complex velocity of a particle of the wake curve, and this by our definition is also

$$\dot{\bar{z}}_s - i\omega z_s + \bar{w},$$

so that the integral is $\rho\int_{z_0}^{z_1} m_s(\dot{\bar{z}}_s - i\omega z_s + \bar{w})\,ds$

This of course could have been written down from our result in paragraph 3 for a single vortex The intrinsic terms are here negligible, as they represent a sum of the type $\Sigma d_s(m_s ds)^2$ Thus the extra terms in the expression for the force due to the presence of the wake combined with the circulation are

$$-\rho\int_{z_0}^{z_1}m_s[\bar{z}_s - ca_1 - c\bar{a}_0 e^{i\zeta_s} + ca_0 e^{i\zeta_s}]\,ds,$$

$$-\rho\int_{z_0}^{z_1}m_s[\dot{\bar{z}}_s + ic\bar{a}_0(\zeta_s e^{i\zeta_s} - \dot{\bar{\zeta}}_s e^{i\zeta_s})]\,ds,$$

$$+\rho\bar{w}K - \rho i c\omega K\bar{a}_1 + \rho i\omega\int_{z_0}^{z_1}m_s[z_s - c\bar{a}_0 e^{i\bar{\zeta}_s} - c\bar{a}_1 + c\bar{a}_0 e^{i\zeta_s}]\,ds,$$

$$+\rho\int_{z_0}^{z_1}m_s(\dot{\bar{z}}_s - i\omega z_s + \bar{w})\,ds,$$

and remembering that $\quad K+\int_{z_0}^{z_1} m_s\,ds = 0,$

these terms reduce to

$$-\rho\int_{z_0}^{z_1} m_s[\bar{z}_s - c\bar{a}_1 - c\bar{a}_0 e^{i\zeta_s} + c\bar{a}_0 e^{i\zeta_s}]\,ds,$$

$$-\rho i c\bar{a}_0\int_{z_0}^{z_1} m_s[\zeta_s e^{i\zeta_s} - \dot{\zeta}_s e^{i\bar{\zeta}_s}]\,ds + \rho i\omega c a_0\int_{z_0}^{z_1} m_s(e^{i\zeta_s} - e^{i\bar{\zeta}_s})\,ds$$

Thus, finally, including the terms found in "I", we have the complete expression for the components of the force on the cylinder

$$\begin{aligned} Y+iX = {} & -\pi i\rho\{Ac^2\bar{w} - 2\bar{B}c^2 w - i\omega c^3\bar{D}\} \\ & -\pi\rho\omega\{Ac^2\bar{w} - 2\bar{B}c^2 w - i\omega c^3\bar{D}\} - \rho\dot{K}c\{a_0 + \sum_{n=2}^{\infty}\bar{a}_n\} \\ & -\rho\int_{z_0}^{z_1} m_s[\bar{z}_s - c\bar{a}_1 - c a_0 e^{i\zeta_s} + c\bar{a}_0 e^{i\zeta_s}]\,ds \\ & -\rho i c\bar{a}_0\int_{z_0}^{z_1} m_s[\zeta_s e^{i\zeta_s} - \dot{\zeta}_s e^{i\bar{\zeta}_s}]\,ds + \rho i\omega c a_0\int_{z_0}^{z_1} m_s(e^{i\zeta_s} - e^{i\bar{\zeta}_s})\,ds \end{aligned}$$

4 5—The couple can be evaluated in the same way The extra acceleration term due to a single vortex $m_s ds$ at the point z_s will be

$$-\tfrac{1}{2}\rho\int_c\left\{\frac{\partial\Omega}{\partial\zeta_s}\zeta_s + \overline{\frac{\partial\Omega}{\partial\zeta_s}}\dot{\zeta}_s + \frac{\partial\Omega}{\partial\bar{\zeta}_s}\dot{\zeta}_s + \overline{\frac{\partial\Omega}{\partial\bar{\zeta}_s}}\zeta_s + \frac{\partial\Omega}{\partial m_s}m_s + \overline{\frac{\partial\Omega}{\partial m_s}}m_s\right\} z\,d\bar{z}$$

As before this becomes

$$-\frac{\rho m_s ds}{2\pi}\int_0^{2\pi}\left\{\zeta_s\frac{e^{-i(\zeta-\zeta_s)}}{1-e^{-i(\zeta-\zeta_s)}} + \dot{\zeta}_s\frac{e^{i(\zeta-\bar{\zeta}_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\} z\frac{d\bar{z}}{d\zeta}\,d\zeta,$$

$$-\frac{\rho i m_s ds}{2\pi}\int_0^{2\pi}\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\} z\frac{dz}{d\zeta}\,d\zeta$$

For the couple we require the real part which is

$$-\frac{\rho m_s ds}{4\pi}\int_0^{2\pi}\left\{\zeta_s\frac{e^{-i(\zeta-\zeta_s)}}{1-e^{-i(\zeta-\zeta_s)}} + \dot{\zeta}_s\frac{e^{i(\zeta-\bar{\zeta}_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\}\frac{d(z\bar{z})}{d\zeta}\,d\zeta,$$

$$-\frac{\rho i m_s ds}{4\pi}\int_0^{2\pi}\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\}\frac{d(z\bar{z})}{d\zeta}\,d\zeta$$

Now on the cylinder

$$z\bar{z} = c^2\left\{\sum_{n=1}^{\infty} b_n e^{ni\zeta} + \sum_{n=1}^{\infty}\bar{b}_n e^{-ni\zeta}\right\},$$

so that
$$\frac{d(z\bar{z})}{d\zeta} = c^2 i\left\{\sum_{n=1}^{\infty} nb_n e^{ni\zeta} - \sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\zeta}\right\}$$

Also as before

$$\zeta_s \frac{e^{-i(\zeta-\zeta_s)}}{1-e^{-i(\zeta-\zeta_s)}} + \bar{\zeta}_s \frac{e^{i(\zeta-\bar{\zeta}_s)}}{1-e^{i(\zeta-\zeta_s)}} = \zeta_s \sum_{n=1}^{\infty} e^{-ni(\zeta-\zeta_s)} + \bar{\zeta}_s \sum_{n=1}^{\infty} e^{ni(\zeta-\bar{\zeta}_s)}$$

and
$$\log\left\{\frac{1-e^{-i(\zeta-\zeta_s)}}{1-e^{i(\zeta-\zeta_s)}}\right\} = \sum_{n=1}^{\infty} \frac{1}{n}\{e^{ni(\zeta-\bar{\zeta}_s)} - e^{-ni(\zeta-\zeta_s)}\}$$

The integrals will therefore be

$$-\tfrac{1}{2}\rho i c^2 m_s ds\left\{\zeta_s \sum_{n=1}^{\infty} nb_n e^{ni\zeta_s} - \bar{\zeta}_s \sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}_s}\right\},$$

$$-\tfrac{1}{2}\rho c^2 m_s ds \sum_{n=1}^{\infty}\{b_n e^{ni\zeta_s} + \bar{b}_n e^{-ni\bar{\zeta}_s}\},$$

so that for a trail of vortices the couple due to the variations in the positions and strengths of the vortices is

$$-\tfrac{1}{2}\rho i c^2 \int_{z_0}^{z_1} m_s\left\{\zeta_s \sum_{n=1}^{\infty} nb_n e^{ni\zeta_s} - \bar{\zeta}_s \sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}_s}\right\} ds,$$

$$-\tfrac{1}{2}\rho c^2 \int_{z_0}^{z_1} m_s \sum_{n=1}^{\infty}\{b_n e^{ni\zeta_s} + \bar{b}_n e^{-ni\bar{\zeta}_s}\} ds$$

4 6—The steady motion terms in the couple are the real part of

$$\tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z}) z\, d\bar{z} - \tfrac{1}{2}\rho\int_c \left(\bar{w} - i\omega z - \frac{d\Omega}{dz}\right)^2 z\, dz,$$

where
$$\frac{d\Omega}{d\zeta} = \frac{d\Omega_a}{d\zeta} - \frac{K}{2\pi} - \int_{z_0}^{z_1} \frac{m_s}{2\pi}\left\{1 - \frac{\sinh \eta_s}{\cosh \eta_s - \cos(\zeta - \xi_s)}\right\} ds$$

The terms
$$\tfrac{1}{2}\rho\int_c (w + i\omega z)(\bar{w} - i\omega\bar{z}) z\, d\bar{z}$$

and
$$-\tfrac{1}{2}\rho\int_c \left\{(\bar{w} - i\omega\bar{z})^2 - 2(\bar{w} - i\omega\bar{z})\frac{d\Omega_a}{dz}\right\} z\, dz$$

are calculated as in "I" The terms introduced by the circulation and wake are again of two types First we have

$$\rho\int_0^{2\pi} (\bar{w} - i\omega\bar{z})\left[\frac{K}{2\pi} - \int_{z_0}^{z_1} \frac{m_s ds}{2\pi}\left\{\frac{e^{-i(\zeta-\zeta_s)}}{1-e^{-i(\zeta-\zeta_s)}} + \frac{e^{i(\zeta-\bar{\zeta}_s)}}{1-e^{i(\zeta-\bar{\zeta}_s)}}\right\}\right] z\, d\zeta$$

in this the terms in K are

$$\rho c a_1 \bar{w} K - \rho i c^2 b_0 \omega K$$

Next there are the terms

$$-\frac{c\rho\bar{w}}{2\pi}\int_{z_0}^{z_1} m_s ds \int_0^{2\pi}\left\{\sum_{n=1}^{\infty} e^{-ni(\zeta-\zeta_s)} + \sum_{n=1}^{\infty} e^{ni(\zeta-\bar{\zeta}_s)}\right\}\left\{\sum_{n=0}^{\infty} a_n e^{(n-1)i\zeta}\right\} d\zeta,$$

which are equal to

$$-\rho\bar{w}\int_{z_0}^{z_1} m_s\{z_s - ca_0 e^{-i\zeta_s} - ca_1 + ca_0 e^{-i\bar{\zeta}_s}\} ds$$

and then

$$\frac{\rho i \omega c^2}{2\pi}\int_{z_0}^{z_1} m_s ds \int_0^{2\pi}\left\{\sum_{n=1}^{\infty} e^{-ni(\zeta-\zeta_s)} + \sum_{n=1}^{\infty} e^{ni(\zeta-\bar{\zeta}_s)}\right\}\left\{\sum_{n=1}^{\infty} (b_n e^{ni\zeta} + \bar{b}_n e^{-ni\bar{\zeta}})\right\} d\zeta,$$

which is $$\rho i \omega c^2 \int_{z_0}^{z_1} m_s \left[\sum_{n=1}^{\infty} (b_n e^{ni\zeta_s} + \bar{b}_n e^{-ni\bar{\zeta}_s})\right] ds$$

The real parts of these are

$$\tfrac{1}{2}\rho c K(a_1\bar{w} + \bar{a}_1 w) - \tfrac{1}{2}\rho i c^2 \omega K(b_0 - \bar{b}_0),$$

$$-\tfrac{1}{2}\rho\bar{w}\int_{z_0}^{z_1} m_s\{z_s - ca_0 e^{-i\zeta_s} - ca_1 + ca_0 e^{-i\bar{\zeta}_s}\} ds,$$

$$-\tfrac{1}{2}\rho w\int_{z_0}^{z_1} m_s\{z_s - c\bar{a}_0 e^{i\zeta_s} - c\bar{a}_1 + c\bar{a}_0 e^{i\bar{\zeta}_s}\} ds$$

Finally, there is $$-\tfrac{1}{2}\rho\int_c \left(\frac{d\Omega}{dz}\right)^2 z\, dz,$$

which is $$\tfrac{1}{2}\rho\int_{z_0}^{z_1}\left\{\left(\frac{d\Omega_1}{dz}\right)^2 - \left(\frac{d\Omega_2}{dz}\right)^2\right\} z_s \frac{dz_s}{ds} ds,$$

and as in the previous case of the force this is

$$\rho\int_{z_0}^{z_1} m_s(\dot{\bar{z}}_s - i\omega z_s + \bar{w})\, z_s ds,$$

the real part of which is

$$\tfrac{1}{2}\rho\int_{z_0}^{z_1} m_s\left\{\frac{d}{dt}(z_s\bar{z}_s) + (\bar{w}z_s + \omega z_s)\right\} ds$$

Thus the extra terms in the expression for the couple due to the presence of the wake combined with the circulation are

$$-\tfrac{1}{2}\rho i c^2\int_{z_0}^{z_1} m_s\left\{\zeta_s\sum_{n=1}^{\infty} nb_n e^{ni\zeta_s}-\bar{\zeta}_s\sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}_s}\right\}ds$$

$$-\tfrac{1}{2}\rho c^2\int_{z_0}^{z_1} m_s\left\{\sum_{n=1}^{\infty}(b_n e^{ni\zeta_s}+\bar{b}_n e^{-ni\bar{\zeta}_s})\right\}ds+\tfrac{1}{2}\rho c K(a_1\bar{w}+a_1 w)$$

$$-\tfrac{1}{2}\rho\bar{w}\int_{z_0}^{z_1} m_s(z_s-ca_0e^{-i\zeta_s}-ca_1+ca_0e^{-i\zeta_s})\,ds$$

$$-\tfrac{1}{2}\rho w\int_{z_0}^{z_1} m_s(\bar{z}_s-c\bar{a}_0e^{i\bar{\zeta}_s}-ca_1+ca_0e^{i\zeta_s})\,ds$$

$$+\tfrac{1}{2}\rho\int_{z_0}^{z_1} m_s\left\{\frac{d}{dt}(z_s\bar{z}_s)+(\bar{w}z_s+w\bar{z}_s)\right\}ds$$

Thus finally, including the terms found in "I", we have for the couple the expression

$$-\pi\rho Cc^4\omega-\tfrac{1}{2}\pi\rho c^3 i\{D\bar{w}-\bar{D}w\}-\tfrac{1}{2}\pi\rho i c^2 w\bar{w}(A-\bar{A})$$

$$-\tfrac{1}{2}K\rho c^2\sum_{n=1}^{\infty}(b_r+\bar{b}_r)-\tfrac{1}{2}\rho c^2\int_{z_0}^{z_1} m_s\sum_{n=1}^{\infty}\{b_n e^{ni\zeta_s}+\bar{b}_n e^{-ni\bar{\zeta}_s}\}\,ds$$

$$+\pi\rho i c^2\{\bar{B}w^2-B\bar{w}^2\}-\tfrac{1}{2}\pi\rho\omega c^3\{\bar{D}w+D\bar{w}\}$$

$$+\tfrac{1}{2}\rho\int_{z_0}^{z_1} m_s\left[\frac{d}{dt}(z_s\bar{z}_s)-ic^2\left\{\zeta_s\sum_{n=1}^{\infty} nb_n e^{ni\zeta_s}-\bar{\zeta}_s\sum_{n=1}^{\infty} n\bar{b}_n e^{-ni\bar{\zeta}_s}\right\}\right]ds$$

$$+\tfrac{1}{2}\rho ca_0\bar{w}\int_{z_0}^{z_1} m_s\{e^{-i\zeta_s}-e^{-i\bar{\zeta}_s}\}+\tfrac{1}{2}\rho ca_0 w\int_{z_0}^{z_1} m_s\{e^{i\zeta_s}-e^{i\bar{\zeta}_s}\}\,ds$$

These formulae are all quite general Beyond the underlying fundamental hydrodynamical assumptions and those relating to the shape and constitution of the wake, no restrictions have been imposed in the derivation They therefore represent the forces and couples on the general cylinder under a generality of conditions which seems the utmost attainable in this theory all that therefore remains is their interpretation

5—THE THIN AEROFOILS WITH THICKNESS AND CAMBER

5 1—The general formulae we have now obtained are rather too complex to convey directly any idea of their physical significance To examine their bearing on the practical problems of aerodynamics it is necessary to con-

sider their form in certain special cases approximating more or less to those realized in practice Of course, as the fundamental hydrodynamical assumptions represent an ideal theoretical limit never really attained in practical problems, we can only expect our formulae to give, in general, rough estimates indicating the order of magnitude of the effects and their general tendency, so there is little except academic interest in developing them beyond a mere first approximation

As the possibility of obtaining direct solutions has in the past been very limited, the discussions of the general effects we have now under investigation have been made mainly under the assumption that the aerofoil is a flat plate This, of course, is a first approximation to the practical shape, but it is one that, in reality, fails to account for the detailed characteristics of the air flow with which we are now more directly concerned These characteristics are rather the result of two further properties of the aerofoil—its blunt fore-end and accompanying thickness, and its camber—and it is necessary therefore in a complete theory to take these into account Various attempts have been made in the past to obtain estimates of the influence of these factors on the behaviour of the aerofoil, but the indirect and rather piecemeal method of procedure adopted is, as we shall see by comparison with results obtained by our direct method, not entirely free from criticism

In the next section of this paper it is proposed therefore to obtain and discuss the first approximation to the previous general results when they are applied to a nearly flat plate, but with a small fore-end thickness and camber The two effects will be dealt with separately, the effect of thickness in a symmetrical aerofoil, and the effect of camber in a thin aerofoil in the shape of a circular arc of small angle The main (long axis) of the section of the aerofoil cylinder will be assumed in both cases to be the axis of x in the co-ordinates on the cross-section plane, so that the thickness and camber are both represented by small variations in the y co-ordinates of the section curve

The previous necessity for deriving results indirectly from the statical case has restricted the derivation of formulae for any but approximately rectilinear motions of the aerofoil, and in a direction to which it can be only slightly inclined While our discussion is not so far restricted in this way, the detailed examination of the effect of the trailing wake becomes very difficult in any other case We shall therefore follow the usual practice and assume for the moment that the aerofoil is moving with a finite velocity V in the direction inclined to the axis of x at a small (variable) angle of attack α We also assume that the aerofoil has a transverse velocity U over and above this main forward velocity, and also an angular velocity $\omega(=\dot{\alpha})$, but that these are so small that we can neglect their squares and products Thus in our final formulae we shall write

$$u = V\cos\alpha, \quad v = V\sin\alpha - U, \quad \omega = \dot{\alpha}$$

Under these circumstances we know from experience that the trailing wake is very approximately straight back along the path of the trailing edge, and as the effect of the wake is of the order of the angle of attack α, the slight inclination of this wake line to the axis of x in our co-ordinate system can also be neglected

5 2—Let us take first the case of the symmetrical aerofoil This is a special case of the thin aerofoil discussed in "I" and is defined by the transformation

$$z = c\sum_{n=0}^{3} a_n e^{ni\zeta},$$

where $\quad a_0 = 1+\epsilon, \quad a_1 = \epsilon, \quad a_2 = 1-\epsilon, \quad a_3 = -\epsilon,$

and ϵ is real and so small that we can neglect its square With these values for the a's we have to a first approximation

$$b_0 = 2, \quad b_1 = \epsilon, \quad b_2 = 1, \quad b_3 = -\epsilon, \quad b_n = 0 \ (n > 3),$$

and also $\quad A = 2, \quad B = 1, \quad C = 2, \quad D = 0$

Our assumptions about the trailing wake mean that it leaves the trailing edge at the point $\zeta = \pi$ and that it lies along the line $\xi = \pi$ Actual distances —from the origin—along this line can be interpreted in terms of η_s by the formula $(\zeta_s = \pi + i\eta_s)$

$$-z_s = s = 2c\cosh\eta_s + 2c\epsilon\sinh\eta_s(1+\cosh\eta_s - \sinh\eta_s)$$

The components of the force on the cylinder due to the fluid pressures in this case are therefore given by

$$\begin{aligned} Y + iX = & -4\pi\rho c^2 v + 4\pi\rho c^2 i\omega v - \dot{K}\rho c(2-\epsilon) \\ & -\rho\int_{z_0}^{z_1} m_s\{s - c\epsilon - 2c(1+\epsilon)\sinh\eta_s\}\,ds \\ & -2\rho c(1+\epsilon)\int_{z_0}^{z_1} m_s\eta_s\cosh\eta_s\,ds + 2\rho i\omega c(1+\epsilon)\int_{z_0}^{z_1} m_s\sinh\eta_s\,ds \end{aligned}$$

To a first approximation, since m_s itself is small, we can neglect in the integrals the thickness of the aerofoil, that is, we assume, with Wagner, that the aerofoil is a flat plate, we have then in this term

$$s = 2c\cosh\eta_s, \quad z_0 = 2c,$$

so that

$$Y + iX = -4\pi\rho c^2 v + 4\pi\rho c^2 i\omega v - \dot{K}\rho c(2-\epsilon)$$
$$-\rho\int_{2c}^{s_1} m_s\{s - \sqrt{(s^2-4c^2)}\}\,ds$$
$$-2\rho c\int_{2c}^{s_1} m_s\{\eta_s \cosh\eta_s + i\omega\sinh\eta_s\}\,ds$$

The couple is similarly

$$-2\pi\rho c^4\omega - 4\pi\rho c^2 uv - \dot{K}\rho c^2$$
$$-\rho\int_{2c}^{s_1} m_s\{\tfrac{1}{2}s^2 - c^2 - \tfrac{1}{2}s\sqrt{(s^2-4c^2)}\}\,ds$$
$$+\rho c^2\int_{2c}^{s_1} m_s\eta_s\cosh 2\eta_s\,ds - 2\rho cu\int_{2c}^{s_1} m_s\sinh\eta_s\,ds,$$

where we have again neglected ϵ in the integrals

Of course the total circulation determined by the constant K is small of the same order as the total wake strength which it balances, so that the term $\dot{K}\rho c\epsilon$ in the expression for the force is also negligible Thus our first conclusion is that to a first approximation a (symmetrical) small thickness has no effect whatever on either the force or couple exerted on the cylinder

5 3—Before discussing the other terms let us now obtain the corresponding results relating to the aerofoil with small camber and no thickness The aerofoil in the shape of a circular arc of small camber proportional to ϵ is defined by the transformation (Durand 1935, p 299)

$$z = c\sum_{n=0}^{3} a_n e^{ni\zeta},$$

where $a_0 = 1,\quad a_1 = i\epsilon,\quad a_2 = 1,\quad a_3 = -i\epsilon,\quad a_n = 0\ (n > 3)$

and, to a first approximation,

$$b_0 = 2,\quad b_1 = -i\epsilon,\quad b_2 = 1,\quad b_3 = -i\epsilon,\quad b_n = 0\ (n>3)$$

and $A = 2,\quad B = 1,\quad C = 2,\quad D = -2i\epsilon,$

but a distinction here arises in that the wake does not leave the cylinder at the point $\zeta_t = \pi$ but rather at the point near this where

$$\zeta_t = \pi - \epsilon$$

This does not, of course, affect the approximate form of our results, for the complex force and couple are, with the same assumptions about the motion,

$$Y+\iota X = -4\pi\rho c^2 v - 2\pi\iota\rho c^3\epsilon\omega - \dot{K}\rho c(2+\iota\epsilon) + 4\pi\rho c^2\iota\omega v$$
$$-2\pi\rho c^3\epsilon\omega^2 - \rho\int_{2c}^{z_1} m_s\{s-\sqrt{(s^2-4c^2)}\}\,ds$$
$$-2\rho c\int_{2c}^{z_1} m_s\{\eta_s\cosh\eta_s+\iota\omega\sinh\eta_s\}\,ds,$$

and the couple Γ,

$$-2\pi\rho c^2\epsilon u - 2\pi\rho c^4\omega - \dot{K}\rho c^2 - 4\pi\rho c^2 uv - 2\pi\rho c^3\epsilon\omega v$$
$$-\rho\int_{2c}^{z_1} m_s\{\tfrac{1}{2}s^2-c^2-\tfrac{1}{2}s\sqrt{(s^2-4c^2)}\}\,ds$$
$$+\rho c^2\int_{2c}^{z_1} m_s\eta_s\cosh 2\eta_s\,ds - 2\rho cu\int_{2c}^{z_1} m_s\sinh\eta_s\,ds$$

These results again show that to a first approximation, when we neglect the products of the small quantities ϵ, ω and ω, a small (symmetrical) camber has no effect on the lift force or drag, although it affects the couple when there is acceleration This agrees with the results obtained by Burgers in Durand (1935, p 301)

To compare the two it is perhaps better if we interpret our results, which are now practically identical in the two cases, in terms of the U, V, and α of Burgers' work Writing as stated above

$$u = V\cos\alpha,\quad v = V\sin\alpha-U,\quad \omega=\dot{\alpha},$$

we get

$$Y+\iota X = -4\pi\rho c^2 V\alpha - 4\pi\rho c^2 V\omega - 2\dot{K}\rho c + 4\pi\rho c^2\iota\omega(V\alpha-U)$$
$$-\rho\int_{2c}^{z_1} m_s\{s-\sqrt{(s^2-4c^2)}\}\,ds - 2\rho c\int_{2c}^{z_1} m_s\{\eta_s\cosh\eta_s+\iota\omega\sinh\eta_s\}\,ds$$

and

$$\Gamma = -2\pi\rho c^3\epsilon V - 2\pi\rho c^4\omega - \dot{K}\rho c^2 - 4\pi\rho c^2 V(V\alpha-U)$$
$$-\rho\int_{2c}^{z_1} m_s\{\tfrac{1}{2}s^2-c^2-\tfrac{1}{2}s\sqrt{(s^2-4c^2)}\}\,ds$$
$$+\rho c^2\int_{2c}^{z_1} m_s\eta_s\cosh 2\eta_s\,ds - 2\rho cV\int_{2c}^{z_1} m_s\sinh\eta_s\,ds,$$

where the term in ϵ in the couple is missing in the case of the symmetrical aerofoil Burgers gives the result for the lift in a form which apparently shows an effect of camber, but his notation is slightly different from ours, his α being our $\alpha+\epsilon$

5 4—The important difference between the results we have obtained and those given by Burgers is in the effect of the wake Basing himself on the result obtained in the statical theory, Burgers obtains in the lift force the term ρKV for the component due to the circulation and vorticity In our formula this component appears as

$$-\rho\int_{2c}^{z_1} m_s \eta_s \cosh \eta_s ds,$$

or if we use $V_s = \eta_s \sinh \eta_s$ as the velocity of the wake particles relative to, and away from, the trailing edge, as

$$-\rho\int_{2c}^{z_1} m_s V_s \coth \eta_s ds$$

It is, however, an added assumption in this work that the vorticity in the wake, once it is created and leaves the trailing edge, remains stationary in the fluid This amounts to assuming that $V_s = V$ or that the force component we are discussing is

$$-\rho V\int_{2c}^{z_1} m_s \coth \eta_s ds$$

But as $\coth \eta_s > 1$ for all positive values of η_s, we have

$$\left|\int_{2c}^{z_1} m_s \coth \eta_s ds\right| > \left|\int_{2c}^{z_1} m_s ds\right| = K,$$

so that now the lift force is actually greater than ρKV In other words the wake vorticity with the same resultant circulation is more effective than simple circulation round the cylinder in producing lift (m_s is of course of the opposite sign to K and the negative sign merely balances this)

The effectiveness of the wake circulation may of course be reduced if the particles—as seems probable—do not have the full relative velocity V assumed in this form of the lift term, that is, if instead of remaining actually at rest in the liquid they moved towards the aerofoil If they moved in the opposite direction their effectiveness would, of course, be increased

Although the effects are negligible in the final approximations, our formulae differ also in one other important aspect from those usually given The complex force due to the wake circulation has, when there is no angular velocity, no appreciable component parallel to the motion In other words the lifting effect of the circulation is produced without any appreciable corresponding drag This distinction between the effect of the wake and the effect of a simple circulation, where this component would be represented by a term $\rho K(V\alpha - U)$, is lost sight of in the statical treatment, where the effects of the wake and the effects of the circulation are separately handled

Our conclusion, and with this we shall be content for the present, is that with the exception of this influence on the lift, of the circulation or the wake, our general discussion confirms all the results obtained by previous methods, but it has the advantage over those methods in that it derives all the results in one direct discussion, and gives precise information as to the nature of the approximations involved at each stage In a further note it is hoped to be able to develop further aspects of these and similar problems to a somewhat greater degree of approximation At present the discussion is being carried out in detail for the Joukowski aerofoils for which the general formulae assume simple finite forms

SUMMARY

The discussion of our solution of the problem of a general cylinder moving in liquids is generalized to include the effect of a vortex-sheet wake extending backwards in the liquid from a trailing edge Assuming the wake to be of any shape and to have associated with it a total circulation which is equal and opposite to that of the main circulation round the cylinder, general formulae are derived for the potential function of the fluid motion and the resultant of the fluid pressures on the cylinder boundary

As steps in the analysis, the solutions of the problem of this general cylinder at rest (i) in a uniform stream, (ii) in a stream due to an external line vortex parallel to its axis, are solved in detail

The general formulae are then examined in their approximate form under conditions analogous to those assumed for the particular problems whose solutions are already known, with the object (i) of determining the first order effects of camber and thickness and (ii) of verifying the approximate theoretical results now used in practice The result that the camber and thickness have no effect on the force (lift or drag), and only a small effect on the couple when there is acceleration, is verified but in a form which shows the precise conditions governing the degree of approximation involved

Our result for the lift arising from the circulations, however, differs from that usually given It is shown that under the usual conditions assumed for the wake, the lift force produced is greater than that produced by the corresponding simple circulation, and that this greater lift is attained without the drag which accompanies it in the simple circulation

REFERENCES

Bickley, W G 1929 *Philos Trans* A, **228**, 235–74
Durand, W F 1935 "Aerodynamic Theory," **2** Berlin
Wagner, H 1925 *Z angew Math Mech* **5**, 17–35

The crystal structure of pentaerythritol tetracetate

By T H Goodwin, *John Harling Fellow, University of Manchester,*
and R Hardy

(*Communicated by W L Bragg, F R S —Received* 18 *November* 1937)

I—Introduction

Although the technique for investigating organic crystals by X-rays is now fairly well developed, there are still many important types of compound of which no representative has yet been examined In particular no Fourier analysis of an ester has ever been published and no reliable information is available as to the lengths of C—O and C=O bonds in such substances The investigation of pentaerythritol tetracetate recorded here was made partly to supply this deficiency and partly to develop the study of pentaerythritol (Llewellyn, Cox and Goodwin 1937) A further incentive lay in the doubt which existed at one time regarding the distribution of the valencies of the central carbon atom in the tetracetate, this is discussed in § III

The results of the investigation are presented at once, the methods by which they were deduced being described in the last section

II—The structure

1—*Molecular symmetry—Description of structure*

Pentaerythritol tetracetate, $C(CH_2O\ CO\ CH_3)_4$, crystallizes in the tetragonal space group $P4_2/n$ The unit cell contains two molecules which may be regarded as mirror images They each have fourfold alternating symmetry, and their central carbon atoms at (000) and $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$ must have their valency bonds directed from the centre to the apices of a tetragonal bisphenoid This bisphenoid is found to be the regular tetrahedron required by chemical theory The eight $—CH_2O\ CO\ CH_3$ groups are all equivalent and are related to each other by the symmetry operations of the space group

Table I gives the co-ordinates of the atoms in one of these side chains Apart from the keto oxygen they have a zigzag arrangement and lie approximately in the crystallographic (520) plane, while the carbonyl bond, which is also in this plane, is parallel to the c-axis (fig 1) Since the average direction of the $\diagdown_{\mathrm{C}}\diagup^{\mathrm{C}}\diagdown_{\mathrm{O}}\diagup^{\mathrm{C}}\diagdown_{\mathrm{C}}$ chains is perpendicular to this axis, the molecules may be described as having their mean planes coincident with {002} but with the atoms of the $—CH_2O\ CO\ CH_3$ groups in the {520} planes

Table I

Atom	x/a	x(A)	y/a	y(A)	z/c	z(A)
C	0	0	0	0	0	0
C_1	0 095	1 14	0 040	0 48	0 161	0 89
O_1	0 181	2 17	0 070	0 84	0	0
C_2	0 270	3 24	0 120	1 44	0 125	0 67
O_2	0 280	3 36	0 120	1 44	0 367	2 02
C_3	0 366	4 40	0 133	1 60	0 950	5 22

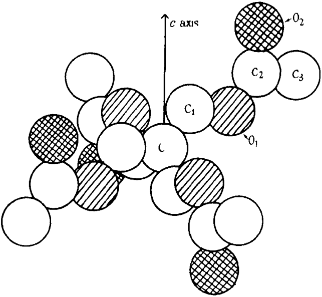


Fig 1—Diagram of pentaerythritol tetracetate molecule $C(CH_2O\ CO\ CH_3)_4$, hydrogens omitted, showing fourfold alternating symmetry of molecule and planar conformation of acetate $—O\ CO\ CH_3$ group

2—*Comparison with pentaerythritol*

The present structure shows a close resemblance to that of pentaerythritol, $C(CH_2OH)_4$ (Llewellyn, Cox and Goodwin 1937) In both, the first side-chain carbon and the adjacent oxygen have z-co-ordinates of 0 89 A and approximately 0 0 A respectively In both, all or nearly all the atoms lie close to the {002} planes, and each has a good cleavage parallel to c{001} though, since the molecules of the ester are not bound together by hydroxyl bonds to form a layer lattice as in the parent alcohol, the former has a much less perfect cleavage, the crystals being considerably softer and having lower density and melting point

3—*Interatomic and intermolecular distances—Valency angles*

The interatomic constants are given in Table II The bond lengths are probably reliable to ± 0 03 A The C—C distances of 1 52 A are in good agreement with those previously measured, viz 1 541 in diamond (Bragg and Bragg 1913), 1 54 ± 0 03 in metaldehyde (Pauling and Carpenter 1936), 1 53 in the normal paraffin series (Müller 1928) and 1 50 ± 0 03 in pentaerythritol Few aliphatic C—O links have been measured, but the most reliable are 1 43 ± 0 03 in metaldehyde, 1 46 ± 0 03 in pentaerythritol and 1 30 in oxalic acid dihydrate (Robertson 1936) The link in metaldehyde may be regarded either as an ether or as a cyclic acetal bond, and so resembles the present C—O links more closely than the others The lengths of 1 41 ± 0 03 found for pentaerythritol tetracetate agree very well with this value, which, it has been suggested (Crowfoot 1936), probably represents the pure single link The double bond between carbon and oxygen has never previously been measured for an ester or similar compound Oxalic acid dihydrate is quoted by Robertson as having a C=O distance of 1 24 He also gives (1934) a length of 1 14 approx in benzoquinone, and Wyckoff (1932, 1934) records 1 25 A for urea However, the present work has included several Patterson and Fourier syntheses, and it is difficult to see that the bond can differ by more than ± 0 03 from 1 33 A

TABLE II

Links	Angles
$C—C_1 = 1\ 52$ A	$C_1CC_1' = 108\frac{1}{2}°$
$C_1—O_1 = 1\ 41$ A	$CC_1O_1 = 105°$
$O_1—C_2 = 1\ 41$ A	$C_1O_1C_2 = 111°$
$C_2—C_3 = 1\ 52$ A	$O_1C_2O_2 = 124°$
$C_2{=}O_2 = 1\ 33$ A	$O_2C_2C_3 = 124\frac{3}{4}°$
	$O_1C_2C_3 = 108\frac{1}{2}°$

The valency angles are probably reliable to ± 3°, and three of the four angles which are approximately tetrahedral do not differ by more than 1½° from 109½°, two of these three are carbon valency angles and the third an oxygen angle which agrees well with Pauling and Carpenter's angle for metaldehyde of 109½ ± 3° It must be admitted that the fourth angle seems rather low, though it is only 4½° less than 109½° The angles between the keto link and the single links in the acetate group are 124° and 124¾° They have never been measured before but are of the order anticipated and agree well with the value of 125½° for benzoquinone The sum of the valency angles round C_2 is 357¼°, and so the acetate group must be regarded as planar within the limits of experimental error This is in accordance with stereochemical theory, but the suggestion that the angle between the single bonds on a doubly bound carbon is greater than 109½° does not appear to be correct

Quite good values are also obtained for intermolecular distances, that between neighbouring keto oxygens being 3 45 A and between adjacent methyl groups 3 52 A

4—*Chemical crystallography of the acetate group*

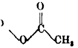
It has been shown above that in accordance with chemical theory the acetate group  has a planar configuration Another interesting feature is that the length of the *c* axis in pentaerythritol tetracetate (5 50 A) is almost identical with the figures found for the shortest axes of a number of acetates of carbohydrates Table III summarizes these data, which are taken from the work of Leuck and Mark (1934) and of Cox and Goodwin (unpublished)

TABLE III

Compound	Axis	Compound	Axis
Glucose pentacetate	5 65 A	Xylose tetracetate	5 72 A
Cellobiose octacetate	5 7 A	Pentaerythritol tetracetate	5 50 A
Cellotriose undecacetate	5 7 A		

Now a study of molecular models shows that these short axes can only be the dimensions of the molecule perpendicular to the sugar ring (cf Cox, Goodwin and Wagstaff 1935) It seems clear, therefore, that in the sugar acetates the —O CO CH_3 groups probably have the same conformation as in pentaerythritol tetracetate, the atoms lying in planes perpendicular to the pyranose ring with the mean direction of each chain parallel to it The detailed examination of these complex compounds should, consequently, be considerably facilitated by the knowledge of the configuration of the acetate group provided by this work

III—DETERMINATION OF THE STRUCTURE

1—*Previous work*

The earliest crystallographic work on pentaerythritol tetracetate was that of Miss Knaggs (1923), but the first X-ray measurements were carried out by Gerstäcker, Moller and Reis (1928), who determined the dimensions of the tetragonal cell and recorded the space group as $P4/n$ (C_{4h}^3). They also concluded that, contrary to chemical ideas, the molecule must have simple fourfold symmetry with a pyramidal distribution of the valencies of the central carbon atom.

In reply to this, Miss Knaggs (1929) showed that the space group had been incorrectly deduced, that it was in fact, $P4_2/n$ (C_{4h}^4), and that the molecular symmetry is fourfold alternating, the valencies of the central carbon being directed from the centre to the apices of a tetragonal bisphenoid. Whether the chemical theory that this bisphenoid is a regular tetrahedron is correct can only be definitely ascertained by the determination of the co-ordinates of the atoms adjacent to the central one. Although Miss Knaggs did not attempt this, she did suggest a tentative structure for the molecule, but as this structure differed radically from that found by us and as it was not supposed to be final no good purpose would be served by discussing it here.

2—*Crystallography, cell dimensions and space group*

The crystals used in this investigation were prepared from pentaerythritol and recrystallized from aqueous alcohol. They formed needles elongated parallel to *c* with usually only the $m\{110\}$ faces present. They were generally terminated by cleavage faces parallel to the base $\{001\}$. The X-ray measurements were carried out by means of rotation and oscillation photographs using a camera of 5·0 cm. radius with Cu $K\alpha$ radiation. They confirmed the cell dimensions recorded by previous workers, viz. $a = 12{\cdot}00$ Å and $c = 5{\cdot}50_6$ Å. These lead to an axial ratio of $a:c = 1:0{\cdot}458$ which is in good agreement with the crystallographic value (Knaggs 1923)

$$a:c = 1:0{\cdot}324\,\sqrt{2} = 1:0{\cdot}458$$

when this refers to the true unit cell. Analysis of oscillations about the [001] and [110] axes showed that all general planes $\{hkl\}$ were present but that $\{hk0\}$ planes with $h+k$ odd were halved. The only order of $\{001\}$ observed was (002). The space group is therefore $P4_2/n$ (C_{4h}^4).

Since the density is 1·273 g./c.c. two molecules of $C(CH_2O\cdot CO\cdot CH_3)_4$ are in the unit cell and each must have fourfold alternating symmetry, the eight $-CH_2O\cdot CO\cdot CH_3$ chains being regarded as equivalent asymmetric groups If one of the central carbon atoms is placed at (000), the other is at $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, and the positions of eight equivalent points are

$$(xyz),\quad (\bar{y}x\bar{z}),\quad (\bar{x}\bar{y}z),\quad (y\bar{x}\bar{z}),\quad (\tfrac{1}{2}+x,\ \tfrac{1}{2}+y,\ \tfrac{1}{2}-z),\quad (\tfrac{1}{2}-y,\ \tfrac{1}{2}+x,\ \tfrac{1}{2}+z),$$

$$(\tfrac{1}{2}-x,\ \tfrac{1}{2}-y,\ \tfrac{1}{2}-z),\quad (\tfrac{1}{2}+y,\ \tfrac{1}{2}-x,\ \tfrac{1}{2}+z)$$

As the hydrogen atoms are too light to be resolved, fifteen parameters need evaluating, therefore, to locate all the heavy atoms in the elementary cell

3—*Intensities*

To determine these parameters photometric measurements of the intensities of as many reflections as possible were made Two methods were adopted to overcome the difficulty that the intensity-blackening curve is not linear over the whole working range In the first, each photograph had recorded on it a calibration strip so that, after photometering any spot, the successive readings could be calibrated in terms of the intensity producing them In the second only the linear portion of the blackening-intensity curve was used, the photographs being taken in duplicate so that strong reflections could be measured on weaker photographs and vice versa The photometry was carried out by means of a Cambridge Microphotometer, the computations of integrated intensity being made directly on a Sundstrand Adding Machine Observed structure factors, F_{obs}, were then calculated by use of the formula $F_{\text{obs}} = KI_m^{\frac{1}{2}}c^{\frac{1}{2}}L^{-\frac{1}{2}}$, where I_m is the integrated intensity as measured on the film—only relative values have been employed as absolute intensities have not been found essential—c is the Cox and Shaw (1930) correction factor, L is the "polarization" factor and K a factor of proportionality

4—*Molecular model*

Although this substance was the first ester to be analysed by X-rays, it was not anticipated that it would show any wide divergence of its interatomic distances from the values of 1 54 and approximately 1 47 A usually assumed for the C—C and C—O links On the other hand the C=O bond should be considerably shorter than 1 47 It was therefore convenient to build a molecular model using spheres of radius 1 5 cm for the atoms C, C_1, O_1 and C_2, fastening them together with concealed wooden "valencies" placed at the tetrahedral angle Actually the carbonyl oxygen O_2 was a sphere of the same size but fastened to C_2 by double spring bonds, in measuring intermolecular distances it was necessary to allow an appropriately closer approach to the centre of the atom The methyl groups were represented by larger spheres of radius 3 5 cm cut away on one side to permit the centre to approach to 3 0 cm (1 5 A) from C_2 In this way methyl radicals could be allowed to "touch", since the nearest distance of approach of carbon atoms in different molecules is never less than 3 5 A The assumption of tetrahedral valency angles for carbon and oxygen was only made as a starting point for preliminary calculations of structure amplitudes and for convenience in studying the space-filling properties of the model

5—*Patterson projection—Space filling—Optics*

At the outset it seemed probable that the molecule would lie in the unit cell with its arms inclined to the *a*{100} faces, for the (110), (310) and (130) reflections are relatively strong while (200) is weak In addition it was found that space-filling criteria were satisfied better with such a molecular arrangement The observations were further confirmed by a Patterson (1935) synthesis (fig 2) using all the {*hk*0} intensities Owing to the large number of interatomic vectors which are only approximately the same in magnitude and direction the maxima are not very clearly revealed, but the development of ridges along the traces of the {520} planes shows that, as a first approximation, the —CH_2O CO CH_3 groups lie in these planes

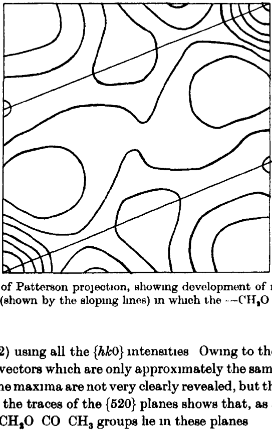


FIG 2—Portion of Patterson projection, showing development of ridges parallel to the {520} planes (shown by the sloping lines) in which the —CH_2O CO CH_3 groups are found

Information as to the configuration of the molecule is also given by the refractive indices $\epsilon = 1\ 483$ and $\omega = 1\ 433$ (Knaggs 1923) The positive birefringence ($\epsilon - \omega = 0\ 050$) is rather large for such a compound and suggests that the keto bonds are parallel to the c-axis, since negative or very low positive double refraction would otherwise be expected Now, if the atoms of each side chain form a flat zigzag, either the C=O or the C—CH_3 bond will be parallel to the principal axis The optical properties, therefore, suggest the former to be the case

6—*Structure factor graphs*

With the information and assumptions already indicated structure factors, F, were calculated for various planes using the formulae

$$F^2 = (\Sigma fA)^2 + (\Sigma fB)^2,$$

$$A = 8\cos^2 2\pi\ \frac{h+k+l}{4}\ \cos\pi[\overline{h-k}\,x + \overline{h+k}\,y]\ \cos\pi[\overline{h+k}\,x - \overline{h-k}\,y]\ \cos 2\pi lz,$$

$$B = -8\sin^2 2\pi\ \frac{h+k+l}{4}\ \sin\pi[\overline{h-k}\,x + \overline{h+k}\,y]\ \sin\pi[\overline{h+k}\,x - \overline{h-k}\,y]\ \sin 2\pi lz$$

The evaluation of these expressions was carried out by means of the contoured graphs of structure factor described by Bragg and Lipson (1936) Since this crystal belongs to the plane group P4 the graphs are typified by fig 6 of their paper As used by us these curves only evaluated the expression

$$A' = 4\cos\pi[\overline{h-k}\,x + \overline{h+k}\,y]\ \cos\pi[\overline{h+k}\,x - \overline{h-k}\,y],$$

but by a change of origin a similar expression

$$B' = 4\sin\pi[\overline{h-k}\,x + \overline{h+k}\,y]\ \sin\pi[\overline{h+k}\,x - \overline{h-k}\,y]$$

could be read off To compute any special or general $\{hkl\}$ structure amplitude it was then only necessary to multiply A' by $2f\cos 2\pi lz$ or B' by $-2f\sin 2\pi lz$ according as $h+k+l$ is even or odd and to sum for all non-equivalent atoms in the cell

The actual curves employed by us were drawn to cover $\frac{1}{2} \times \frac{1}{2}$ of the cell on a sheet 20 × 20 in By plotting atomic positions on a sheet of tracing cloth which could be placed over each graph in turn, the contribution of each atom to any geometrical structure factor could be read off

7—{*hk*0} *structure amplitudes and Fourier projections*

In spite of the information and assumptions indicated earlier, considerable latitude still existed in the choice of atomic co-ordinates It was found impossible at first to obtain satisfactory agreement between observed and calculated values of F_{hk0}, but it was realized that certain planes always gave structure factors with the same signs, notably (110), (310), (130) and (330) A Fourier projection was therefore constructed using these four terms only and served to restrict the possible atomic positions Various new pairs of co-ordinates were then employed in computing the structure factors afresh and when more planes were found to have constant signs they too were incorporated in a second synthesis and the process repeated In this way the atomic co-ordinates were gradually refined and the agreement between observed and calculated F's improved until the projection shown in fig 3 was obtained

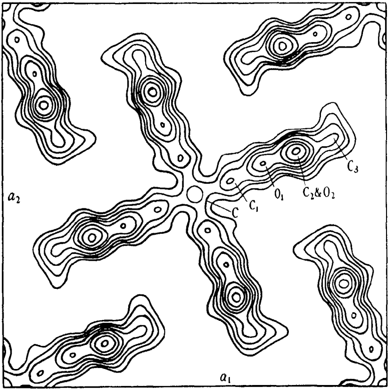


FIG 3—Last Fourier projection on (001), showing arrangement of arms of molecule in {520} planes

It may be mentioned here that in all computations of $\{hk0\}$ structure factors the effect of interchanging O_2 and C_3 was tried to determine definitely whether C_2—O_2 or C_2—C_3 is parallel to the c-axis Invariably more satisfactory results were obtained in the former case

8—{*hkl*} *structure factors—Three-dimensional Fourier syntheses*

In calculating the values of F_{hkl} it was assumed in the first instance that C_1 would lie at a height of 1 54 cos ($\frac{1}{2} \times 109\frac{1}{2}°$) = 0 89 A, and that O_1 and C_3 would be in the same plane as the central carbon atom i e $z = 0$ The heights of C_2 and O_2 were found from a Patterson synthesis (Harker 1936) along the line (x_{c_2}, y_{c_2}, z), this synthesis is shown in fig 4 Using these z-co-ordinates

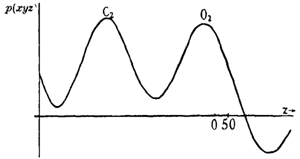


FIG 4—One dimensional Patterson synthesis parallel to c axis through C_2 and O_2 to determine z co ordinates of C_2 and O_2

with the values of x and y found from fig 3, structure amplitudes were computed for a large number of $\{hkl\}$ planes and those which showed good agreement with the experimental figures were incorporated in one-dimensional Fourier syntheses through C_2O_2 and O_1, from which more reliable values were obtained for the z-co ordinates The structure factors were then recalculated and many more included in new syntheses both along special lines and over various planes at different heights in the cell The latter helped to determine the x and y co ordinates of various atoms, particularly C_2 and O_2, more accurately than was possible with the projections on $\{001\}$ For these syntheses the density of scattering matter at any point is given, in arbitrary units since only relative intensities were used, by

$$\rho_{xyz} = \sum_{-h}^{h} \sum_{-k}^{k} \sum_{-l}^{l} F_{hkl} \cos\{2\pi(hx + ky + lz) - \alpha\},$$

where α, the phrase angle of any reflection, is given by

$$\alpha = \tan^{-1}\frac{\Sigma fB}{\Sigma fA}, \text{ and, for this space group, } = 0,\ \pi/2,\ \pi,\ 3\pi/2$$

The computation of these syntheses is best effected by the separate treatment of terms with $h+k+l$ odd and even as implied by Mrs Lonsdale (1936) and was carried out by an extension of the method of Lipson and Beevers (1936)

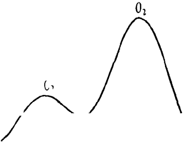
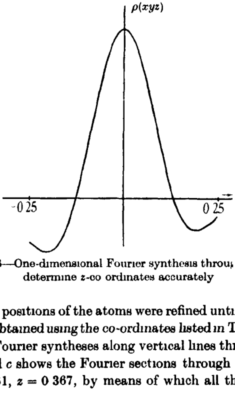


FIG 5—Fourier synthesis along line C_2O_2 to determine
z co ordinates more accurately than from fig 4



IG 6—One-dimensional Fourier synthesis throu
determine *z*-co ordinates accurately

In this way the positions of the atoms were refined until the results shown in Table IV were obtained using the co-ordinates listed in Table I Figs 5 and 6 show the final Fourier syntheses along vertical lines through C_2O_2 and O_1 and fig 7 *a*, *b* and *c* shows the Fourier sections through the cell at heights of $z = 0$, $z = 0\ 161$, $z = 0\ 367$, by means of which all the atoms are com-

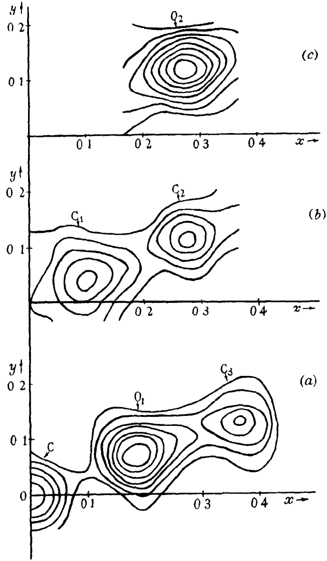


FIG 7—Fourier sections through unit cells at different heights

(a) at $z = 0$, showing central C, O_1 and C_3,
(b) at $z = 0\ 161$, showing C_1 and C_2,
(c) at $z = 0\ 367$, showing O_2 (keto oxygen)

pletely resolved The negative ordinates in figs 5 and 6 are due to the omission of the $\{hk0\}$ terms as they make contributions which are independent of z

TABLE IV

hkl	F_{calc}	F_{obs}	α	hkl	F_{calc}	F_{obs}	α
110	37	42	0	530	21	13	0
020	12	10	180	251	25	23	180
011	71	35	0	521	51	42	0
111	43	29	270	042	18	13	180
220	8	9	180	060	17	15	180
021	39	30	90	412	3	17	270
121	12	26	180	441	10	21	270
211	37	33	0	332	7	< 11	180
130	58	44	180	242	20	26	180
310	118	92	0	422	34	37	0
221	2	14	270	351	41	33	270
131	5	10	270	531	20	22	270
311	11	26	270	260	20	10	180
040	38	47	180	620	62	38	0
231	21	22	180	013	14	15	0
321	9	14	180	113	13	13	270
330	46	42	180	512	39	42	0
002	56	39	0	261	21	18	90
240	21	26	0	023	16	22	90
420	55	58	0	123	37	38	180
012	18	18	270	522	3	15	270
112	6	9	0	550	8	< 9	0
041	32	32	270	170	8	9	0
141	37	29	180	710	14	21	180
411	24	27	0	460	21	12	0
331	22	26	90	640	18	14	180
122	3	< 10	270	361	39	24	0
212	4	11	90	223	5	13	270
241	7	10	90	033	11	16	180
421	52	39	270	352	12	16	0
222	13	19	180	532	16	19	0
150	10	9	0	551	15	12	90
510	3	< 6	180	641	16	16	90
032	32	35	90	370	3	< 11	180
132	19	19	180	730	42	30	0
312	12	12	0	323	5	21	0
431	30	24	0	612	38	16	270
051	18	16	0	271	12	*ca* 12	0
151	17	17	90	043	24	18	270
232	22	25	270	333	22	17	90
322	19	22	270	262	29	28	0
440	26	22	180	622	14	< 14	0
350	6	< 10	0	413	18	17	0

The planes in this table are listed in decreasing order of spacing

TABLE IV—*continued*

hkl	F_{calc}	F_{obs}	α	*hkl*	F_{calc}	F_{obs}	α
080	16	9	0	930	50	16	0
561	25	16	180	314	28	12	180
731	24	19	270	363	34	24	0
820	46	14	180	044	0	<12	—
253	11	12	0	273	26	16	0
523	37	27	0	244	36	18	180
750	42	25	180	950	18	10	0
372	44	9	0	514	47	23	0
004	14	<12	180	444	39	18	0
661	23	14	270	2, 10, 0	17	10	180
163	24	19	0	10, 2, 0	23	10	0
024	15	12	0	4, 10, 0	13	<10	180
453	23	24	0	10, 4, 0	53	15	0
190	31	11	180	11, 1, 0	40	12	0
910	20	7	180	264	40	15	0
390	25	<10	0	941	16	14	180

The authors desire to express their warmest thanks to Professor W L Bragg, F R S , for his continued interest and helpful discussions and to the University of Manchester for the John Harling Fellowship held by one of us (T H G) while this work was carried out They also thank Professor I M Heilbron, F R S , for his interest and Dr S Peat for preparing the material

IV—SUMMARY

The crystal structure of pentaerythritol tetracetate has been determined The molecule has fourfold alternating symmetry, the valencies of the central carbon atom being inclined to each other at the tetrahedral angle The remaining valency angles are listed, with the interatomic distances, in Table II They show that the acetate group is planar, as anticipated by chemical theory and that the length of the C=O bond is 1 33 A

REFERENCES

Bragg, W H and Bragg, W L 1913 *Proc Roy Soc* A, **89**, 277
Bragg, W L and Lipson, H 1936 *Z Kristallogr* **95**, 323
Cox, E G , Goodwin, T H and Wagstaff, A E 1935 *J Chem Soc* p 1495
Cox, E G and Shaw, W F B 1930 *Proc Roy Soc* A, **127**, 71
Crowfoot, D 1936 *Rep Progr Chem* **33**, 220
Gerstäcker, A , Moller, H and Reis, A 1928 *Z Kristallogr* **66**, 355
Harker, D 1936 *J Chem Phys* **4**, 381
Knaggs, I E 1923 *J Chem Soc* **123**, 77

Knaggs, I E 1929 *Proc Roy Soc* A, **122**, 69
Leuck, G J and Mark, H 1934 *J Amer Chem Soc* **56**, 1960
Lipson, H and Beevers, C A 1936 *Proc Phys Soc* **48**, 772
Llewellyn, F J , Cox, E G and Goodwin T H 1937 *J Chem Soc* p 883
Lonsdale, K 1936 "Structure Factor Tables" London Bell
Müller, A 1928 *Proc Roy Soc* A, **120**, 437
Patterson, A L 1935 *Z Kristallogr* **90**, 517
Pauling, L and Carpenter, D C 1936 *J Amer Chem Soc* **58**, 1274
Robertson, J M 1934 *Proc Roy Soc* A, **146**, 106
— 1936 *J Chem Soc* p 1817
Wyckoff, R W G 1932 *Z Kristallogr* **81**, 102
— 1934 *Z Kristallogr* **89**, 462

The electronic structure of some polyenes and aromatic molecules

IV—The nature of the links of certain free radicals

By C A Coulson

University Chemical Laboratory and Trinity College, Cambridge

(*Communicated by J E Lennard-Jones, F R S —
Received* 24 *November* 1937)

1 Introduction

In the first two papers of this series (1937*a*, *b*), referred to as I, II, Lennard-Jones has developed a method for investigating the lengths and energies of the links in some unsaturated molecules For this purpose he used the method of molecular orbitals In paper III (1937), Penney obtained similar results, using the electron-pair methods of resonance, developed by Pauling and others It is the purpose of the present paper to extend the calculations to chain molecules and radicals in which the number of carbon atoms is odd, and in which, therefore, there is one electron which does not form a bond, in the usual picture of the chemist We shall use the method of molecular orbitals, and this work may be said to be a continuation of I and II The writer would like to thank Professor Lennard-Jones for suggesting this work, and for the opportunity of discussing it with him during the calculations

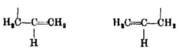
In general these free radicals with "trivalent" carbon are not stable, and tend to form dimers, but there are certain of them which do exist either as stable substances or in dissociative equilibrium with their dimers Huckel (1935) has discussed these radicals, on the earlier form of the theory in which all the links were assumed equal and no allowance was made for their compression His work needs to be extended because there is no reason why the links should be all equal, and in fact, the bond diagrams of the chemist lead one to suspect otherwise, and to believe that there may be one of the carbon atoms (the one on which the unpaired electron is to be found) different from the others (for which all the electrons are paired) On the molecular orbital theory, in which each electron is supposed free to move throughout the whole molecule in an averaged potential field, it is not so easy to see at once in what way the presence of the odd electron will alter the arrangement of the links So the first question that we shall ask will be whether in chain molecules, such as $C_{2n+1}H_{2n+3}$, there is one carbon atom occupying an essentially different situation from all the others We shall then compute the resonance energy

The details of the calculations, the setting up of the secular determinant, the inclusion of the energy of compression of the C—C single links and the minimizing of the total mobile energy $\mathcal{F}$ with respect to the lengths of the individual links are supposed known, they are described in detail in I The values for the constants s, d, k_s and k_d are taken to be the same as in I and II

2—THE CHAIN MOLECULES $C_{2n+1}H_{2n+3}$

(i) C_3H_5

The simplest of these chain molecules is the radical C_3H_5, and it will be convenient to discuss it in some detail before working through the general case This molecule would naturally be described by either of the two bond diagrams of fig 1



$H_2\dot{C}—\underset{|}{\underset{H}{C}}=CH_2 \qquad H_2C=\underset{|}{\underset{H}{C}}—\dot{C}H_2$

FIG 1

On the wave-mechanical picture, there will be resonance between the two patterns, and this will result, as calculation shows, in this case, in a perfectly symmetrical structure in which the central atom is equally spaced from the two end radicals The calculation is as follows

Call the lengths of the two links x_1 and x_2, so that the resonance integrals are β_1 and β_2. Then the secular determinant is

$$\begin{vmatrix} \epsilon & \beta_1 & 0 \\ \beta_1 & \epsilon & \beta_2 \\ 0 & \beta_2 & \epsilon \end{vmatrix} = 0,$$

i.e.
$$\epsilon = 0 \quad \text{or} \quad \epsilon^2 = \beta_1^2 + \beta_2^2 \tag{1}$$

Three mobile electrons are to be put into this framework, and their electronic energy $\mathscr{E}$ may be written

$$\mathscr{E} = 2\{\beta_1^2 + \beta_2^2\}^{\frac{1}{2}}$$

The compressional energy $\mathscr{V}$ is

$$\mathscr{V} = k_s(x_1 - s)^2 + k_s(x_2 - s)^2,$$

and so the total orbital energy is

$$\mathscr{F} = 2\{\beta_1^2 + \beta_2^2\}^{\frac{1}{2}} + k_s(x_1 - s)^2 + k_s(x_2 - s)^2 \tag{2}$$

The defining equations are $\dfrac{\partial \mathscr{F}}{\partial x_1} = \dfrac{\partial \mathscr{F}}{\partial x_2} = 0$.

Just as in I, these equations simplify, and the result is

$$x_1 = s - (s-d) \Big/ \left[1 + \frac{k_s}{k_d}\left\{\left(1 + \frac{\beta_2^2}{\beta_1^2}\right)^{\frac{1}{2}} - 1\right\}\right], \tag{3}$$

$$x_2 = s - (s-d) \Big/ \left[1 + \frac{k_s}{k_d}\left\{\left(1 + \frac{\beta_1^2}{\beta_2^2}\right)^{\frac{1}{2}} - 1\right\}\right] \tag{4}$$

The solution of these equations shows that in equilibrium $x_1 = x_2 = 1.36_6$ Å. (All distances will be given in Ångstrom units.) It is easily verified that this is indeed a minimum for $\mathscr{F}$, since, with these values,

$$\frac{\partial^2 \mathscr{F}}{\partial x_1^2} = \frac{\partial^2 \mathscr{F}}{\partial x_2^2} > 0 \quad \text{and} \quad \frac{\partial^2 \mathscr{F}}{\partial x_1^2}\,\frac{\partial^2 \mathscr{F}}{\partial x_2^2} > \left(\frac{\partial^2 \mathscr{F}}{\partial x_1\,\partial x_2}\right)^2$$

We conclude therefore that the effect of resonance between the two structures of fig. 1 is to give a perfectly symmetrical structure, in which the unpaired electron must be considered as moving throughout the whole of the molecule. It is interesting to note how much charge may be thought of as residing on each nucleus. If we call the three carbon nuclei 1, 2 and 3, in order, and write, for the molecular orbitals,

$$\Psi = a_1\psi_1 + a_2\psi_2 + a_3\psi_3,$$

where ψ_1 is the wave-function of a π-electron on nucleus 1, then a rough description of the charge on each nucleus is obtained if we say that the contributions from the orbital Ψ are in the ratio a_1^2 a_2^2 a_3^2 respectively For the lowest orbital a_1 a_2 $a_3 = 1$ $\sqrt{2}$ 1 so that the charge distribution due to the two electrons in this orbital is 0 5, 1 0 and 0 5 respectively For the upper orbital (singly occupied) a_1 a_2 $a_3 = 1$ 0 -1, so that the charge distribution for this orbital is 0 5, 0, 0 5 Thus the total charge distribution for the three π-electrons, being the sum of the two effects above, is 1, 1, 1, indicating that the charge does not tend to accumulate on the central carbon atom nor on the outer two carbon atoms

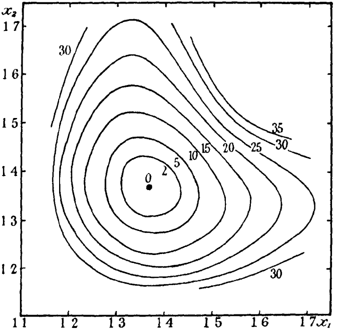


FIG 2—Energy contours for C_3H_5 (Distances in A, energies in kcal)

The energies of the occupied orbitals are 46 4 kcal and 0 The energy of compression is 21 7 kcal , so that the total mobile energy $\mathscr{F}$ is -71 7 kcal Taking the energy of a simple double bond as 55 7 kcal , we find a resonance energy of 15 4 kcal

It is easy to see why this radical will try to form a dimer For if two such radicals come together to form a chain molecule

$$H_2C{=}CH{-}CH_2{-}CH_2{-}CH{=}CH_2,$$

there will be a gain of energy equal to that of a single C—C bond (72 6 kcal) and a loss, relative to the two isolated radicals, of twice the resonance energy of one radical (i e 2×15 4 kcal) There is thus a net gain of energy of about 42 kcal

Equation (2) enables $\mathscr{F}$ to be computed for any assigned values of x_1 and x_2, and energy contours can be drawn, which show how the energy of the molecule depends upon x_1 and x_2 Such contours, at intervals of 5 kcal , are shown in fig 2 For values of x outside the range shown, it is not justifiable to replace the σ-bonds between the carbon atoms by a parabolic field $k_s(x-s)^2$, but a full Morse-curve should be used The energy surface of fig 2 could be used to compute the vibrational frequencies, for which the normal co-ordinates are clearly $x_1 \pm x_2$

(ii) *Chain molecules* $C_{2n+1}H_{2n+3}$

The general radical of this type is $C_{2n+1}H_{2n+3}$ and can be treated similarly, if we suppose that the links alternate in length, as in the ordinary conjugated chains of even order, then there will be two structures between which resonance may be expected In each there will be n single and n double bonds, and the unpaired electron will be on that carbon atom at the end of the chain which is terminated by a single bond The two structures will be similar except that, starting from one end, the first link is single in the one case and double in the other In this connexion, "single" and "double" refer to the customary bond representation the actual lengths of the links will certainly not be precisely those corresponding to a pure single or double bond Let us call the lengths alternately x_1 and x_2 with exchange integrals β_1 and β_2, then the secular determinant is

$$\Delta \equiv \begin{vmatrix} \epsilon, & \beta_1, & 0, & , & , & , & 0, & 0 \\ \beta_1, & \epsilon, & \beta_2, & , & , & , & 0, & 0 \\ 0, & \beta_2, & \epsilon, & \beta_1, & , & , & 0, & 0 \\ & & & & & & & \\ 0, & 0, & 0, & , & , & \beta_1, & \epsilon, & \beta_2 \\ 0, & 0, & 0, & , & , & 0, & \beta_2, & \epsilon \end{vmatrix} = 0 \qquad (5)$$

In this determinant there are $2n+1$ rows and columns Now Δ is just the determinant R_{n+1} discussed in the appendix to this paper, and its value is there obtained In fact,

$$\Delta = \beta_1^n \beta_2^n \epsilon \sin^{n+1}\phi/\sin\phi,$$

where

$$\epsilon^2 = \beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\cos\phi \qquad (6)$$

The roots are given by

$$\epsilon_0 = 0$$

and $$\epsilon_r = \{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos r\pi/(n+1)\}^{\frac{1}{2}}, \quad r = 1, 2, 3, \quad n \tag{7}$$

The levels corresponding to $r = 1, 2, 3, \quad n$ (with positive root for ϵ_r) will each be completely filled, this accounts for $2n$ mobile electrons, and the $2n+1$th will partly fill the level ϵ_0 Thus

$$\mathscr{E} = 2\sum_1^n \{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos r\pi/(n+1)\}^{\frac{1}{2}}$$

But $$\mathscr{V} = nk_s(x_1 - s)^2 + nk_s(x_2 - s)^2,$$

and $$\mathscr{F} = \mathscr{E} + \mathscr{V} \tag{8}$$

The defining equations, $\dfrac{\partial \mathscr{F}}{\partial x_1} = \dfrac{\partial \mathscr{F}}{\partial x_2} = 0$ now yield the relations

$$x_1 = s - (s-d)\Big/\left\{1 + \frac{k_s}{k_d}\left(\frac{n}{S_1} - 1\right)\right\}, \tag{9a}$$

$$x_2 = s - (s-d)\Big/\left\{1 + \frac{k_s}{k_d}\left(\frac{n}{S_2} - 1\right)\right\}, \tag{9b}$$

where $$S_1 = \sum_1^n \left(\beta_1 + \beta_2 \cos\frac{r\pi}{n+1}\right)\Big/\left\{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos\frac{r\pi}{n+1}\right\}^{\frac{1}{2}}, \tag{10a}$$

$$S_2 = \sum_1^n \left(\beta_2 + \beta_1 \cos\frac{r\pi}{n+1}\right)\Big/\left\{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos\frac{r\pi}{n+1}\right\}^{\frac{1}{2}} \tag{10b}$$

These equations are satisfied by $x_1 = x_2$, to show that this is an absolute minimum we need to show that $\dfrac{\partial^2 \mathscr{F}}{\partial x_1^2}$ and $\dfrac{\partial^2 \mathscr{F}}{\partial x_2^2}$ are greater than 0, and that $\dfrac{\partial^2 \mathscr{F}}{\partial x_1^2}\dfrac{\partial^2 \mathscr{F}}{\partial x_2^2} > \left(\dfrac{\partial^2 \mathscr{F}}{\partial x_1 \partial x_2}\right)^2$ We have seen that this condition is satisfied in the case of $n = 1$, and it can be shown to be true also for $n = 2$ and as n tends to infinity, in other cases it is easy to verify that it is satisfied, and we conclude therefore that radicals of type $C_{2n+1}H_{2n+3}$ do not show alternating lengths of links analogous to the ordinary conjugated systems

We can therefore suppose that each link is of equivalent length (this statement needs to be modified when we allow for end-effects in the next section) In this case the secular determinant (5) simplifies When $\beta_1 = \beta_2 = \beta$, it becomes

$$P_{2n+1}(\beta) = 0, \qquad \text{(see appendix for the value of } P_n)$$

i.e. $$\sin (2n+2)\theta/\sin\theta = 0,$$

where $$\epsilon = 2\beta\cos\theta$$

The roots are $$\epsilon_r = 2\beta\cos r\pi/(2n+2),$$

where $$r = 1, 2, 3, \ldots, 2n+1 \tag{11}$$

The levels $r = 2n+1, 2n, \ldots, n+2$ are doubly occupied and the level $r = n+1$ is singly occupied (its energy $\epsilon_{n+1} = 0$). Thus

$$\mathscr{E} = 4\beta\sum_{1}^{n}\cos r\pi/(2n+2) = 2\beta[\cot\pi/(4n+4) - 1] \tag{12}$$

Now $$\mathscr{V} = 2nk_s(x-s)^2 \tag{13}$$

Therefore $$\mathscr{F} = 2nk_s(x-s)^2 + 2\beta[\cot\pi/(4n+4) - 1] \tag{14}$$

The defining equation now gives

$$x = s - (s-d)\bigg/\left[1 + \frac{k_s}{k_d}\,\frac{(2n+1)\tan\pi/(4n+4) - 1}{1 - \tan\pi/(4n+4)}\right] \tag{15}$$

This last equation enables the equilibrium distance to be calculated for any value of n. Equation (11) then gives the energies of the various orbitals. It will be noted that as n tends to infinity,

$$x \to s - (s-d)\bigg/\left\{1 + \frac{k_s}{k_d}\left(\frac{\pi}{2} - 1\right)\right\} \tag{16}$$

This is in entire agreement with the corresponding results for an even number of carbon atoms as in I (p. 293, equation (56)). As $n\to\infty$ the energy per carbon atom tends to

$$\frac{1}{2n+1}\{2\beta[\cot\pi/(4n+4) - 1] + 2nk_s(x-s)^2\},$$

and in the limit this becomes $4\beta/\pi + k_s(x-s)^2$, i.e. 31.0 kcal. This latter value will be the same for chains containing an even number of carbon atoms. Numerical results for a number of different values of n are shown in Table I. In this table, some of the results for the molecules discussed in I are included for comparison and then the value for the length given in the table is the mean length of the links. Attention is also drawn to the fact that the resonance energy is measured from a "zero" of n double bonds both in the cases $C_{2n}H_{2n+2}$ and $C_{2n+1}H_{2n+3}$.

TABLE I—THE LENGTHS OF THE LINKS IN CHAIN MOLECULES AND RADICALS C_xH_{x+2}

	C_3H_5	C_4H_6	C_5H_7	C_6H_8	C_7H_9	C_8H_{10}	C_9H_{11}	$C_{10}H_{12}$	$C_\infty H_\infty$
Length of end links	1.36_6	—	1.35_2	—	1.34_9	—	1.34_8	—	1.33_7
Length of other links	—		1.39_1	—	1.38_5	—	1.38_3	—	1.37_7
Mean length	1.36_6	1.36_5	1.37_2	1.37_0	1.37_3	1.37_3	1.37_4	1.37_4	1.37_7
Resonance (kcal.)	15.4	5.0	25.5	10.8	33.0	17.1	40.2	22.3	—
Mean length (all links assumed equal)	1.36_6	1.35_9	1.37_0	1.36_8	1.37_2	1.37_1	1.37_3	1.37_2	1.37_7
Gain in energy by allowing end links to differ from the rest	0.0	—	0.76	—	0.82	—	0.88	—	1.65
Total energy content due to mobile electrons	658.8	897.6	1084.6	1292	1481	1687	1877	2081	—

(iii) *End-effects in the radicals* $C_{2n+1}H_{2n+3}$

It has been shown in the preceding section that in the radicals $C_{2n+1}H_{2n+3}$ there is no reason to expect an alternating series of lengths for the various links, and the calculations subsequently described were those in which it was assumed that all the links were of equal length. This, however, is not entirely fair, for, as Penney shows in III, the links at the end may be appreciably shorter than those in the central part of the molecule. Penney concluded that in molecules of type $C_{2n}H_{2n+2}$ the end-effects persisted up to the fourth link from either end. With molecules of type $C_{2n+1}H_{2n+3}$, where the initial tendency to alternate sizes does not exist, we may reasonably expect the end-effects to be rather less important. An approximate result could be obtained by assuming that all the links were of equal length x_2 except the two end ones, of length x_1. With this assumption, the secular determinant becomes

$$\Delta \equiv \begin{vmatrix} \epsilon, & \beta_1, & 0, & , & , & , & , & , & 0 \\ \beta_1, & \epsilon, & \beta_2, & , & , & , & , & , & 0 \\ 0, & \beta_2, & \epsilon, & \beta_2, & , & , & , & , & 0 \\ & & & & & & & & \\ 0, & 0, & 0, & , & , & \beta_2, & \epsilon, & \beta_2, & 0 \\ 0, & 0, & 0, & , & , & 0, & \beta_2, & \epsilon, & \beta_1 \\ 0, & 0, & 0, & , & , & 0, & 0, & \beta_1, & \epsilon \end{vmatrix} = 0 \qquad (17)$$

This determinant may be expanded in powers of β_1. It gives

$$\Delta = \epsilon^2 P_{2n-1} - 2\epsilon\beta_1^2 P_{2n-2} + \beta_1^4 P_{2n-3}$$

If we use the recurrence relations for the P_n, as given in the appendix, this reduces to

$$\Delta = P_{2n+1} - 2(\beta_1^2 - \beta_2^2)\, P_{2n-1} + (\beta_1^2 - \beta_2^2)^2\, P_{2n-3} \tag{18}$$

The roots may be found as a series of approximations, just as in the case of I (equations (34), (35)), we find that

$$\theta_r = \frac{r\pi}{2n+2} - \frac{1}{n+1}\left(\frac{\beta_1^2}{\beta_2^2} - 1\right)\sin\frac{2r\pi}{2n+2} - \frac{n-1}{2(n+1)^2}\left(\frac{\beta_1^2}{\beta_2^2} - 1\right)^2 \sin\frac{4r\pi}{2n+2} \tag{19}$$

Since $(\beta_1^2/\beta_2^2 - 1)$ is small, an expansion in power series converges fairly rapidly With this value of θ_r, we have

$$\begin{aligned} \epsilon_r &= 2\beta_2 \cos\theta_r \quad (r = 1, 2 \ldots 2n+1) \\ &= 2\beta_2 \cos\frac{r\pi}{2n+2} + \frac{\beta_2}{n+1}\left(\frac{\beta_1^2}{\beta_2^2} - 1\right)\left(\cos\frac{r\pi}{2n+2} - \cos\frac{3r\pi}{2n+2}\right) \end{aligned} \tag{20}$$

We have neglected powers of $(\beta_1^2/\beta_2^2 - 1)$ above the first in this approximation There are $2n+1$ mobile electrons to fit into these levels The levels $r = 1, 2 \ldots n$ are doubly filled, and the level $r = n+1$ (for which $\epsilon = 0$ rigorously) is singly occupied Hence

$$\mathscr{E}(x_1, x_2) = 4\beta_2 \sum_1^n \cos\frac{r\pi}{2n+2} + \frac{2\beta_2}{n+1}\left(\frac{\beta_1^2}{\beta_2^2} - 1\right)\sum_1^n\left(\cos\frac{r\pi}{2n+2} - \cos\frac{3r\pi}{2n+2}\right), \tag{21}$$

$$\mathscr{V}(x_1, x_2) = 2k_s(x_1 - s)^2 + (2n-2)\, k_s(x_2 - s)^2, \tag{22}$$

$$\mathscr{F}(x_1, x_2) = \mathscr{E}(x_1, x_2) + \mathscr{V}(x_1, x_2) \tag{23}$$

Now $$\sum_1^n \cos r\pi/(2n+2) = \tfrac{1}{2}\{\cot\pi/(4n+4) - 1\}$$

and $$\sum_1^n \cos 3r\pi/(2n+2) = \tfrac{1}{2}\{-\cot 3\pi/(4n+4) - 1\},$$

so that

$$\mathscr{E}(x_1, x_2) = 2\beta_2\left[\cot\frac{\pi}{4n+4} - 1\right] + \frac{\beta_2}{n+1}\left(\frac{\beta_1^2}{\beta_2^2} - 1\right)\left(\cot\frac{\pi}{4n+4} + \cot\frac{3\pi}{4n+4}\right) \tag{24}$$

The defining equations $\partial\mathscr{F}/\partial x_1 = \partial\mathscr{F}/\partial x_2 = 0$ now yield the results

$$x_1 = s - (s-d)\Big/\left\{1 + \frac{k_s}{k_d}\left(\frac{1}{S_1} - 1\right)\right\}, \tag{25a}$$

$$x_2 = s - (s-d)\Big/\left\{1 + \frac{k_s}{k_d}\left(\frac{1}{S_2} - 1\right)\right\}, \tag{25b}$$

where
$$S_1 = \frac{1}{2n+2}\,\frac{\beta_1}{\beta_2}\left(\cot\frac{\pi}{4n+4} + \cot\frac{3\pi}{4n+4}\right), \tag{26}$$

$$S_2 = \frac{1}{4n-4}\left[2\cot\frac{\pi}{4n+4} - 2 - \frac{1}{n+1}\left(1+\frac{\beta_1^2}{\beta_2^2}\right)\left(\cot\frac{\pi}{4n+4} + \cot\frac{3\pi}{4n+4}\right)\right] \tag{27}$$

These formulae may be applied for various values of n It is interesting to note that as n tends to infinity, $S_2 \to 2/\pi$ and $x_2 \to 1\ 377$, which is the value found for an infinite chain in which all the links are assumed equal In the case of an infinite chain, $S_1 \to 8\beta_1/3\pi\beta_2$, and the equation (25*a*) must be solved by iteration It becomes

$$x_1 = s - (s-d)\Big/\left[1 + \frac{k_s}{k_d}\left(\frac{3\pi\beta_2}{8\beta_1} - 1\right)\right] \tag{28}$$

The result is that the end links are of length 1 337 A Since it can make no appreciable difference in a very long chain whether the number of carbon atoms is odd or even, this calculation must hold for all long chains It is probable that the estimate would need slight revision if we were to allow for the variation of the next links to the end ones as well, but the calculations would be excessively laborious by this method It should be noted that Penney (III, p 321) estimates the length of the end links in long chains to be 1 35 A, so that the agreement is very satisfactory It may be worth noting that the total gain in energy due to this variation of the end links is 1 7 kcal over the value obtained with all the links equal For shorter molecules of this same type, as we shall see below, the gain is rather less (about 0 8 kcal)

(iv) *Some special cases* C_5H_7, C_7H_9 *and* C_9H_{11}

These methods have been applied numerically to study the radicals C_5H_7, C_7H_9 and C_9H_{11} The end links in each case have been assumed to have a length x_1 different from those of the other links x_2, all of which are supposed equal In these calculations it was possible to minimize directly with respect to x_1 and x_2 without using the approximate formulae shown in equation (20) The results, which are shown in Table 1, are therefore independent of the validity of this equation (which only holds when $(1-\beta_1^2/\beta_2^2)^2$ may be neglected) The values for the mobile energy shown in the table may be compared with the corresponding values when all the links are supposed equal, a difference of about 1 kcal is obtained between the two energies The equilibrium values both of x_1 and x_2 get steadily less as the chain increases in length, tending to their asymptotic values 1 337 and 1 377 respectively, but the mean length of the links is remarkably constant The end link always appears to be about 0 04 A shorter than the other links, a value very comparable with that obtained by Penney (III, p 321)

3—SUMMARY

General formulae are given for the lengths of the links of free radicals and molecules of chain-form $C_{2n+1}H_{2n+3}$ It is shown that in these chains the effect of resonance is to remove some of the characteristic properties of alternate single and double bonds Calculations have been made of the energies of these radicals in certain particular cases, and the effect of resonance has been estimated

APPENDIX

THE VALUES OF CERTAIN DETERMINANTS

Some of the determinants evaluated in this appendix have already been calculated by Lennard Jones in I and II, and by Huckel and others But for convenience their values are included with the new ones, the method used by Lennard-Jones is only applicable to cases where there is an even number of rows and columns, the present method applies quite generally

Let $P_n(d)$ be the determinant

$$\begin{vmatrix} \epsilon, & d, & 0, & 0, & \cdots, & \cdots, & \cdots, & 0 \\ d, & \epsilon, & d, & 0, & \cdots, & \cdots, & \cdots, & 0 \\ 0, & d, & \epsilon, & d, & \cdots, & \cdots, & \cdots, & 0 \\ \cdots & & & & & & & \\ 0, & 0, & 0, & 0, & \cdots, & d, & \epsilon, & d \\ 0, & 0, & 0, & 0, & \cdots, & 0, & d, & \epsilon \end{vmatrix} \tag{29}$$

There are n rows and columns, and the only non-vanishing elements are the leading diagonal elements and those immediately bordering this diagonal Expanding in terms of the top row, we find the recurrence relation

$$P_n = \epsilon P_{n-1} - d^2 P_{n-2} \tag{30}$$

The initial values are $P_1 = \epsilon$ and $P_2 = \epsilon^2 - d^2$ (31)

If we write $\mathcal{P} = 1 + P_1 x + P_2 x^2 + \cdots,$ (32)

then the recurrence relation (30), together with the initial values (31), shows that

$$\mathcal{P} = 1/(1 - \epsilon x + d^2 x^2) \tag{33}$$

If, now, we write $\epsilon = 2d\cos\theta$, and put $\mathscr{P}$ into partial fractions, then we may equate coefficients of x^n in $\mathscr{P}$ and find

$$P_n = d^n \sin(n+1)\theta/\sin\theta \tag{34}$$

Thus the roots of $P_n = 0$ are all real and are

$$\epsilon_r = 2d\cos r\pi/(n+1), \quad \text{where } r = 1, 2, \quad n \tag{35}$$

Now let

$$Q_n(d,s) = \begin{vmatrix} \epsilon & d & 0, & 0, & , & , & , & , & , & 0 \\ d, & \epsilon, & s, & 0, & , & , & , & , & , & 0 \\ 0, & s & \epsilon, & d & , & , & , & , & , & 0 \\ 0, & 0, & d, & \epsilon, & s, & , & , & , & , & 0 \\ \\ 0, & 0, & 0, & 0, & , & , & d, & \epsilon, & s, & 0 \\ 0, & 0, & 0, & 0, & , & , & 0, & s, & \epsilon, & d \\ 0, & 0, & 0, & 0, & , & , & 0, & 0, & d, & \epsilon \end{vmatrix}$$

and

$$R_n(d,s) = \begin{vmatrix} \epsilon, & s, & 0, & 0 & , & , & , & , & , & 0 \\ s, & \epsilon, & d, & 0, & , & , & , & , & , & 0 \\ 0, & d, & \epsilon, & s, & , & , & , & , & , & 0 \\ 0, & 0, & s, & \epsilon, & d, & , & , & , & , & 0 \\ \\ 0, & 0, & 0, & 0, & , & , & d, & \epsilon, & s, & 0 \\ 0, & 0, & 0, & 0, & , & , & 0, & s, & \epsilon, & d \\ 0, & 0, & 0, & 0, & , & , & 0, & 0, & d, & \epsilon \end{vmatrix}, \tag{36}$$

so that Q_n has $2n$ rows and columns and R_n has $2n-1$ Expanding in terms of the top rows, we find in the two cases

$$Q_n = \epsilon R_n - d^2 Q_{n-1},$$
$$R_n = \epsilon Q_{n-1} - s^2 R_{n-1} \tag{37}$$

Elimination of R_n, or of Q_n, provides the new equations

$$Q_n + (s^2 + d^2 - \epsilon^2)\,Q_{n-1} + s^2 d^2 Q_{n-2} = 0, \tag{38}$$

$$R_n + (s^2 + d^2 - \epsilon^2)\,R_{n-1} + s^2 d^2 R_{n-2} = 0 \tag{39}$$

The initial values are $\quad Q_0 = 1, \quad Q_1 = \epsilon^2 - d^2,$ (40)

$$R_0 = 0, \quad R_1 = \epsilon, \quad R_2 = \epsilon^3 - \epsilon(s^2 + d^2) \tag{41}$$

If we write
$$\mathcal{Q} = Q_0 + Q_1 x + Q_2 x^2 + \cdots, \tag{42}$$

then it is soon shown that

$$\mathcal{Q} = (1 + s^2 x)/\{1 + (s^2 + d^2 - \epsilon^2)x + s^2 d^2 x^2\} \tag{43}$$

Putting
$$\epsilon^2 = s^2 + d^2 + 2sd\cos\phi \tag{44}$$

and throwing $\mathcal{Q}$ into partial fractions, the result previously obtained by Lennard-Jones is reproduced

$$Q_n = s^n d^{n-1}\{d\sin(n+1)\phi + s\sin n\phi\}/\sin\phi \tag{45}$$

In an exactly similar way, we find that

$$R_n = s^{n-1} d^{n-1} \epsilon \sin n\phi/\sin\phi \tag{46}$$

A check on this working is obtained since, when $s = d$,

$$\epsilon^2 = d^2(2 + 2\cos\phi) = 4d^2\cos^2(\phi/2),$$

and so, comparing with the relation $\epsilon = 2d\cos\theta$, we see that $\phi = 2\theta$ In such a case Q_n does reduce to P_{2n}, and R_n does reduce to P_{2n-1}

For the sake of completeness it is worth including in this section two determinants which, though not used in the calculations of the present paper, would be required in any discussion of the cyclic molecules and radicals

Let
$$L_n(d,s) = \begin{vmatrix} \epsilon, & d, & 0, & 0, & \cdots, & \cdots, & 0, & s \\ d, & \epsilon & s, & 0, & \cdots, & \cdots, & 0, & 0 \\ 0, & s, & \epsilon, & d, & \cdots, & \cdots, & 0, & 0 \\ 0 & 0, & 0, & 0, & \cdots, & s, & \epsilon, & d \\ s, & 0, & 0, & 0, & \cdots, & 0, & d, & \epsilon \end{vmatrix} \tag{47}$$

L_n contains $2n$ rows and columns Expanding in terms of the top row,

$$L_n(d,s) = \epsilon R_n(d,s) - d^2 Q_{n-1}(d,s) - s^2 Q_{n-1}(s,d) - 2d^n s^n \tag{48}$$

If we put in the previously determined values of R and Q, this reduces to

$$L_n(d,s) = -4d^n s^n \sin^2 n\phi/2, \tag{49}$$

where
$$\cos\phi = (\epsilon^2 - s^2 - d^2)/2sd \tag{50}$$

Lastly, let

$$M_n(d) = \begin{vmatrix} \epsilon, & d, & 0, & 0, & , & , & 0, & d \\ d, & \epsilon, & d, & 0, & , & , & 0, & 0 \\ 0, & d, & \epsilon, & d, & , & , & 0, & 0 \\ & & & & & & & \\ 0, & 0, & 0, & 0, & , & d, & \epsilon, & d \\ d, & 0, & 0, & 0, & , & 0, & d, & \epsilon \end{vmatrix} \tag{51}$$

There are n rows and columns in M_n M_{2n} is a particular case of L_n in which $s = d$ But in the case of M_{2n+1} we require special treatment Expanding in terms of the top row, we obtain

$$M_n = \epsilon P_{n-1} - 2d^2 P_{n-2} + (-1)^{n-1} d^n$$

Inserting the values of P_{n-1} and P_{n-2}, we find the result

$$M_{2n} = -4d^{2n} \sin^2 n\theta, \tag{52}$$

$$M_{2n+1} = 4d^{2n+1} \cos^2 \frac{2n+1}{2} \theta, \tag{53}$$

in which, as usual, $\epsilon = 2d \cos\theta$

The value of M_{2n} agrees with that obtained from L_n when $s = d$ It may be added that the expansions of these determinants show at once that all their roots are real, in the cases of Q, R, L and M it also shows that most of their roots are double

REFERENCES

Lennard-Jones 1937*a* *Proc Roy Soc* A, **158**, 280 I

Lennard Jones and Turkevich 1937*b* *Proc Roy Soc* A, **158** 297 II

Penney 1937 *Proc Roy Soc* A, **158** 306, III

Huckel, E 1935 Physical Soc ' Int Conf Phys " p 1 London

Sidgwick 1933 "Covalent Link in Chemistry", p 112 Oxford Univ Press

The electronic structure of some polyenes and aromatic molecules

V—A comparison of molecular orbital and valence bond methods

By G W Wheland*

(*Communicated by J E Lennard-Jones, F R S — Received 24 November 1937*)

In the preceding papers of this series (Lennard Jones, Turkevich and Penney 1937) it has been found that the molecular orbital and valence-bond methods lead to values for the lengths of links in polyenes and other molecules in satisfactory agreement with each other, in fact it may be said that in most of the applications of these methods to problems of molecular structure the two methods are found to agree, in a roughly qualitative way, with each other and with experiment (Wheland 1934) Since the approximations involved in the two cases are of quite different natures, this fact suggests that both treatments are actually rather more reliable than might have been anticipated in view of the mathematical approximations necessary in both methods It is, accordingly of interest to examine in some detail the examples in which definite discrepancies do occur, in order that the factors responsible may be determined In the present paper this will be done for cyclobutadiene, C_4H_4 For purposes of comparison, a brief discussion of benzene will be given as well, in order that the differences between the two molecules may be brought out more clearly

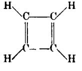
In both the valence-bond and the molecular orbital treatments we shall introduce the usual simplification (Huckel 1931) of neglecting all the orbitals that are symmetrical with respect to reflexion in the plane of the molecule These either belong to the *K* shells and so are not concerned in the binding or else merely contribute a constant additive term to the total energy (in neither case does their neglect seriously affect the following discussion) The problem of C_4H_4 reduces, accordingly, to one of four electrons, which are to be assigned to orbitals that are antisymmetric to such a reflexion In the valence-bond treatment, one electron is assigned, with suitable spin, to each of the four atomic p_z orbitals (the molecule being assumed to lie in the xy plane), and these are then allowed to interact with each other in different ways The binding is, accordingly, considered to be purely covalent The subsequent calculation leads to the result that the ground state of the molecule is a singlet with energy $Q+2\alpha$, whereas the lowest triplet state has the energy Q Q is here the coulomb integral $(abcd \mid H \mid abcd)$, and α is the single exchange integral between two adjacent p_z orbitals,

---

* Foreign Fellow of the John Simon Guggenheim Memorial Foundation, 1936–7, temporarily at the University Chemical Laboratory, Cambridge

$$(abcd \mid H \mid bacd) = (abcd \mid H \mid acbd) = (abcd \mid H \mid abdc) = (abcd \mid H \mid dbca),$$

while a, b, c and d represent the p_z orbitals taken in order round the ring * This result is based upon the assumption that all exchange integrals of the energy except α and all exchange integrals of unity can be neglected If the molecule possessed one of the Kekule-like structures



its energy to the same approximation would be $Q+\alpha$ The difference between these two quantities is termed the resonance energy and is equal simply to α In order to obtain agreement with the observed resonance energies of a number of other hydrocarbons, it has been found necessary to set α equal to about $-1\ 5$ e volts (Pauling and Wheland 1933)

In the molecular orbital treatment, the procedure is to assign the four electrons to various molecular orbitals, which are not confined to individual atoms but are allowed to extend over the entire molecule It is convenient to assume that these can be approximated to by linear combinations of the atomic p_z functions, a, b, c and d, so that

$$\psi_i = k_{ia} a + k_{ib} b + k_{ic} c + k_{id} d,$$

with numerical coefficients k_{ia}, etc When the variational method is applied, the orbitals and their corresponding energies are found to be†

$$\begin{aligned}
\psi_1 &= \tfrac{1}{2}(a+b+c+d)\ w_1 = g+2\beta, \\
\psi_2 &= \tfrac{1}{2}(a+b-c-d)\ w_2 = g, \\
\psi_3 &= \tfrac{1}{2}(a-b-c+d)\ w_3 = g, \\
\psi_4 &= \tfrac{1}{2}(a-b+c-d)\ w_4 = g-2\beta
\end{aligned}$$

---

* These equalities are dependent upon the assumption of a square molecule

† In these and in all the following expressions of the same type, the numerical coefficients multiplying the wave functions are of such magnitude as to normalize them, provided that the various atomic orbitals are themselves normalized and mutually orthogonal The first condition is satisfied in general, but the second is not, so that the resulting functions are not actually normalized It is useful, however, to retain the coefficients

(Since ψ_2 and ψ_3 are degenerate, any two independent linear combinations of them could be used as well.) The coulomb integral, g, is equal here to

$$(a \mid H' \mid a) = (b \mid H' \mid b) = (c \mid H' \mid c) = (d \mid H' \mid d),$$

while the resonance integral β is equal to

$$(a \mid H' \mid b) = (b \mid H' \mid c) = (c \mid H' \mid d) = (d \mid H' \mid a)$$

H' is a one-electron Hamiltonian operator which refers to the self consistent field. In deriving the above equations for the ψ's and the w's, the approximation has been made of neglecting all resonance integrals except β, and of treating a, b, c and d as mutually orthogonal. Two electrons are now assigned, with opposite spins, to the most stable orbital, ψ_1, and the remaining two are assigned to ψ_2 and ψ_3. This can be done in four different ways so as to produce one triplet and three singlet states. To the present approximation, all of these have the same energy, $4g+4\beta$, equal to the sum of the energies of the occupied orbitals (those orbitals being counted twice that are occupied twice). If the molecule possessed one of the Kekulé-like structures, the occupied orbitals would have to be taken as

$$\psi_5 = \frac{1}{\sqrt{2}}(a+b), \quad w_5 = g+\beta$$

and

$$\psi_6 = \frac{1}{\sqrt{2}}(c+d), \quad w_6 = g+\beta,$$

or else as

$$\psi_7 = \frac{1}{\sqrt{2}}(a+d), \quad w_7 = g+\beta$$

and

$$\psi_8 = \frac{1}{\sqrt{2}}(b+c), \quad w_8 = g+\beta$$

In either case, the energy would again be just $4g+4\beta$, so that the calculated resonance energy is zero, in serious disagreement with the result obtained by the valence bond method. One might have expected to find here a resonance energy of about 1.8β, since the results of similar calculations by the two methods for other molecules are generally in entirely satisfactory agreement if β is set equal to about $\frac{5}{9}\alpha$ or about -0.85 e volt (Wheland 1934).

At first sight, the molecular orbital treatment seems to be the better of the two in this case, since cyclobutadiene, unlike the similarly constituted benzene, is apparently too unstable to exist. This fact can be correlated with the calculated lack of resonance energy. The argument is not conclusive, however, because the strain in the four-membered ring must be enormous. A detailed analysis of the problem by Penney (1934) has, in fact, shown that this effect alone is probably sufficient to account for the facts

A more successful line of attack is suggested by the fact that the present sort of disagreement is encountered only in the cases in which the molecular orbital treatment finds the ground state degenerate For cyclobutadiene, it will be remembered, there were three different singlet states with energy $4q+4\beta$ It is natural to suppose that these may interact strongly with each other, so that at least one is greatly stabilized The molecular orbital treatment does not provide a method for estimating the magnitude of the interaction energy, but the following considerations prove illuminating The molecular eigenfunction for the one of the three singlet levels arising from the configuration $\psi_1^2\psi_2^2$ can be expressed, after being made antisymmetric, in the form of a determinant

$$\Phi_1 = \frac{1}{(4!)^{\frac{1}{2}}}\begin{vmatrix} \psi_1(1) & \overline{\psi_1}(1) & \psi_2(1) & \overline{\psi_2}(1) \\ \psi_1(2) & \overline{\psi_1}(2) & \psi_2(2) & \overline{\psi_2}(2) \\ \psi_1(3) & \psi_1(3) & \psi_2(3) & \psi_2(3) \\ \psi_1(4) & \overline{\psi_1}(4) & \psi_2(4) & \psi_2(4) \end{vmatrix}$$

$$= \frac{1}{16(4!)^{\frac{1}{2}}}$$

$$\begin{vmatrix} (a+b+c+d)(1) & (\bar{a}+\bar{b}+c+d)(1) & (a+b-c-d)(1) & (a+\bar{b}-c-\bar{d})(1) \\ (a+b+c+d)(2) & (a+\bar{b}+c+\bar{d})(2) & (a+b-c-d)(2) & (a+\bar{b}-\bar{c}-\bar{d})(2) \\ (a+b+c+d)(3) & (a+\bar{b}+\bar{c}+d)(3) & (a+b-c-d)(3) & (a+\bar{b}-c-\bar{d})(3) \\ (a+b+c+d)(4) & (\bar{a}+\bar{b}+\bar{c}+d)(4) & (a+b-c-d)(4) & (\bar{a}+\bar{b}-\bar{c}-\bar{d})(4) \end{vmatrix}$$

The orbitals over which bars are drawn, as $\overline{\psi_1}(1)$, are associated with negative, and the others with positive spin, so that $\overline{\psi_1}(1)$ represents $\psi_1(1)\beta(1)$, $\psi_1(1)$ represents $\psi_1(1)\alpha(1)$, and so on By adding and subtracting columns, we do not alter the value of the determinant, but we can easily change its form to

$$\Phi_1 = \frac{1}{4(4!)^{\frac{1}{2}}}\begin{vmatrix} (a+b)(1) & (\bar{a}+\bar{b})(1) & (c+d)(1) & (\bar{c}+d)(1) \\ (a+b)(2) & (\bar{a}+\bar{b})(2) & (c+d)(2) & (\bar{c}+\bar{d})(2) \\ (a+b)(3) & (\bar{a}+\bar{b})(3) & (c+d)(3) & (\bar{c}+\bar{d})(3) \\ (a+b)(4) & (a+\bar{b})(4) & (c+d)(4) & (c+\bar{d})(4) \end{vmatrix}$$

$$= \frac{1}{(4!)^{\frac{1}{2}}}\begin{vmatrix} \psi_5(1) & \overline{\psi_5}(1) & \psi_6(1) & \overline{\psi_6}(1) \\ \psi_5(2) & \overline{\psi_5}(2) & \psi_6(2) & \overline{\psi_6}(2) \\ \psi_5(3) & \overline{\psi_5}(3) & \psi_6(3) & \overline{\psi_6}(3) \\ \psi_5(4) & \overline{\psi_5}(4) & \psi_6(4) & \overline{\psi_6}(4) \end{vmatrix}$$

Φ_1, therefore, represents simply the Kekulé-like structure in which the double bonds lie between the atoms *a* and *b* and between *c* and *d* In the same way, the function, Φ_2, arising from the configuration, ψ_1^2, ψ_3^2, is seen to represent the second Kekulé-like structure, since it can be equally well derived from the configuration, ψ_7^2, ψ_8^2 The interaction between Φ_1 and Φ_2 is, therefore, just the resonance between the two Kekulé-like structures The reason for the apparent lack of resonance energy is now clear, the molecular orbital treatment simply neglects it The valence-bond treatment, which deals with it explicitly, takes the calculation to a higher degree of approximation and presumably provides the more trustworthy result

The same considerations also show the reason why Lennard-Jones and Turkevich (1937) found that the bond distances in cyclobutadiene should be equal alternately to those characteristic of pure single and pure double bonds Any treatment, in fact, which neglects the resonance between the two Kekulé-like structures would necessarily lead to that answer In this connexion, it may be mentioned that, when the analogous calculation is carried through by the valence bond method, all carbon-carbon distances are found to be equal to each other and very nearly equal to that in benzene It should not be necessary to give the details of the calculation here, since the procedure is obvious and completely straightforward

One further discrepancy in regard to the energy levels of cyclobutadiene remains to be considered, and for this no completely satisfactory solution seems to be possible To a first approximation, as we have seen, the molecular orbital method finds the ground state of the molecule to be fourfold degenerate and to consist of one triplet and three singlet levels When the next approximation is made of considering the interactions between these levels, the degeneracy is removed, as we have also seen, with the result that some are made more, and some are made less stable The difficulty lies in the fact that when this splitting of the levels is taken into account, the triplet is found to lie lowest The reason for this is the same as in the molecular orbital treatment of oxygen, the orbitals involved are orthogonal to each other, so that their exchange integral is positive In fact, when antisymmetric singlet and triplet functions, Φ_3 and Φ_4, respectively, are set up corresponding to the configuration $\psi_1^2\,\psi_2\,\psi_3$, the following relationships are found

$$W_{\Phi_1+\Phi_2} = W_{\Phi_4} + \left(\phi\phi\left|\frac{1}{r_{12}}\right|\phi\phi\right) + 2\left(\phi\phi\left|\frac{1}{r_{12}}\right|\chi\chi\right) - \left(\phi\chi\left|\frac{1}{r_{12}}\right|\phi\chi\right),$$

$$W_{\Phi_1-\Phi_2} = W_{\Phi_4} + 2\left(\phi\phi\left|\frac{1}{r_{12}}\right|\chi\chi\right),$$

$$W_{\Phi_3} = W_{\Phi_4} + \left(\phi\phi\left|\frac{1}{r_{12}}\right|\phi\phi\right) - \left(\phi\chi\left|\frac{1}{r_{12}}\right|\phi\chi\right)$$

(For reasons of symmetry, the functions, Φ_1 and Φ_2, are not themselves suitable but must be replaced by the combinations $(\Phi_1 \pm \Phi_2)/\sqrt{2}$ This has the effect of introducing automatically the formerly neglected resonance between the two Kekulé-like structures) In these equations,

$$W_{\Phi_1+\Phi_2} = (\Phi_1+\Phi_2 \mid H \mid \Phi_1+\Phi_2)/(\Phi_1+\Phi_2 \mid \Phi_1+\Phi_2),$$

and so on, where H is now the correct Hamiltonian operator and not merely the one referring to the self-consistent field The functions ϕ and χ are defined as

$$\phi = N(a-c),$$
$$\chi = N(b-d),$$

where $N = \dfrac{1}{\sqrt{2-2(a/c)}}$ is a normalization constant It will be seen that the integrals, $\left(\phi\phi\left|\dfrac{1}{r_{12}}\right|\phi\phi\right)$, etc , are all positive, so that the triplet energy, W_{Φ_4}, is indeed the lowest, as stated above (It is necessary also to know something of the relative magnitudes of the integrals, since some occur with negative coefficients The order

$$\left(\phi\phi\left|\frac{1}{r_{12}}\right|\phi\phi\right) > \left(\phi\chi\left|\frac{1}{r_{12}}\right|\phi\chi\right) > \left(\phi\phi\left|\frac{1}{r_{12}}\right|\chi\chi\right)$$

seems reasonable In any case, there is no doubt that the first is the largest of the three—which is sufficient to insure that the triplet state lies lowest) The above equations are exact, since in their derivation no integrals have been neglected and the orbitals, a, b, c and d, have not been assumed orthogonal

This result is directly opposed to that of the valence-bond treatment, which finds the ground state to be a singlet, with the lowest triplet lying higher by -2α or about 3 e volts Since the molecule is not known, the empirical test cannot be applied to decide which answer is correct All that can be done, therefore, is to examine the various wave-functions and determine the relation between them

Of the three singlet states, $(\Phi_1+\Phi_2)/\sqrt{2}$ is without question the highest, as is seen from the above equations This corresponds, in a manner to be described below, to the less stable of the two singlets found by the valence-bond treatment It is not certain which of the two remaining singlet states, $(\Phi_1-\Phi_2)/\sqrt{2}$ and Φ_3, lies lower The first of these corresponds to the state which the valence-bond method makes the ground state, while the second corresponds to a state which the valence-bond method does not take into account at all The function, Φ_4, corresponds to the lowest triplet state found by the valence-bond method

Before discussing the nature, of this correspondence, it will be well to digress for a moment and to consider the relation between the molecular orbital and the valence-bond treatments of hydrogen, H_2 If the atomic $1s$ orbitals about the two nuclei are designated as r and s, respectively, the approximate molecular orbital involved in the normal molecule is $(r+s)/\sqrt{2}$ (There need be only one molecular orbital, since there are only two electrons) The complete wave-function is, accordingly,

$$\Phi = \frac{1}{2\sqrt{2}} \begin{vmatrix} (r+s)(1) & (\bar{r}+\bar{s})(1) \\ (r+s)(2) & (\bar{r}+\bar{s})(2) \end{vmatrix}$$

On expanding the determinant, we find

$$\Phi = \frac{1}{2\sqrt{2}}\{r(1)\,r(2)+s(1)\,s(2)+r(1)\,s(2)+s(1)\,r(2)\}\{\alpha(1)\,\beta(2)-\beta(1)\,\alpha(2)\}$$

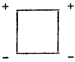
This function can be interpreted in terms of the valence-bond method as representing resonance among the structures

$$r^-s^+, \quad r^+s^- \quad \text{and} \quad r-s$$

The (normalized*) functions corresponding to these structures occur with coefficients of $\frac{1}{2}$, $\frac{1}{2}$ and $1/\sqrt{2}$ respectively

In exactly the same way, the functions, $(\Phi_1+\Phi_2)/\sqrt{2}$, $(\Phi_1-\Phi_2)/\sqrt{2}$, Φ_3 and Φ_4, can be expanded in terms of covalent and ionic structures The calculations are somewhat tedious, but the results are those given in figs 1–4 It is found that the purely covalent parts of the functions, $(\Phi_1+\Phi_2)/\sqrt{2}$, $(\Phi_1-\Phi_2)/\sqrt{2}$ and Φ_4, are identical with the functions given by the valence-bond method for the upper and the lower singlet states and the lowest triplet state, respectively, whereas Φ_3 contains no purely covalent part The difference between the results obtained by the two methods must, therefore, be due to the effect of the partly or wholly ionic structures which occur in the molecular orbital but not in the valence-bond treatment

It is not difficult to see why, if the ionic structures really are as important as the molecular orbital method makes them, the singlet state $(\Phi_1-\Phi_2)/\sqrt{2}$ of cyclobutadiene should be comparatively unstable The totally ionic structures of the type



---

* Cf footnote, p 398

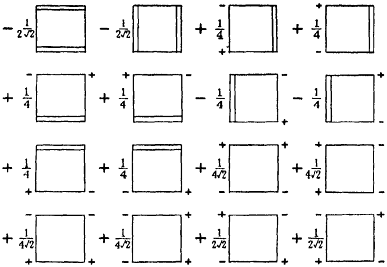


FIG 1 The numbers in front of the structures give the magnitudes of the coefficients of the corresponding functions when these are normalized in the sense of the footnote on p 398 The signs of the functions are determined by the following convention Each function can be expressed as a sum of terms of the form $\pm kW(1)\,X(2)\,Y(3)\,Z(4)$, where k is a positive constant, W is an arbitrary one of the orbital functions, a, b, c or d, multiplied by one of the spin functions, α or β, and similarly for X, Y and Z There will be one such term in which the orbitals appear alphabetically, and the spin functions come in the order $\alpha\beta\alpha\beta$ This term is taken with the positive sign, and then the signs of all the others are determined by the exclusion principle

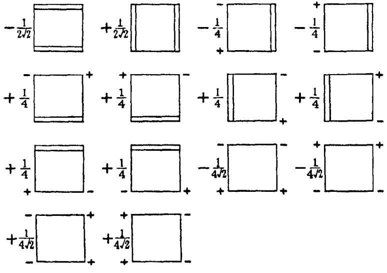


FIG 2 See legend to fig 1

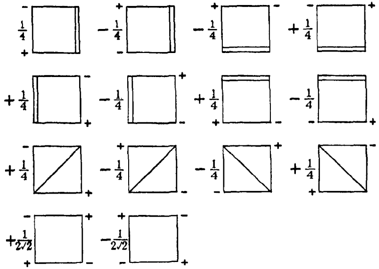


FIG 3 See legend to fig 1

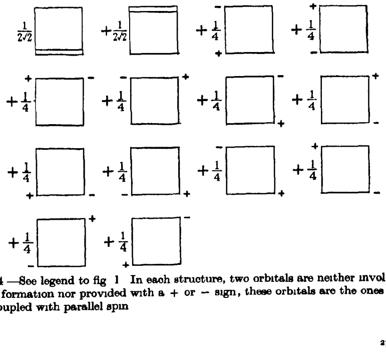
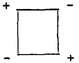
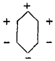


FIG 4 —See legend to fig 1 In each structure, two orbitals are neither involved in bond formation nor provided with a + or − sign, these orbitals are the ones which are coupled with parallel spin

occur with comparatively large coefficients, while the much more reasonable structures of the type



are excluded on grounds of symmetry That this is not a general occurrence is shown by the example of benzene, for which a similar expansion leads to the result shown in fig 5 In this case, the structures



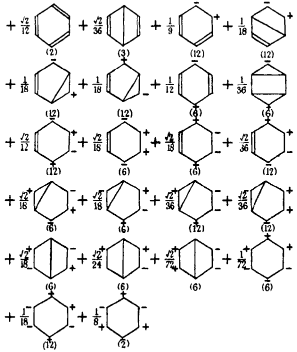
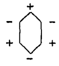


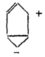
FIG 5—See legend to fig 1 This eigenfunction is formed by assigning two electrons with opposite spins to each of the three molecular orbitals, $\frac{1}{\sqrt{6}}(a+b+c+d+e+f)$, $\frac{1}{2\sqrt{3}}(2a+b-c-2d-e+f)$, and $\frac{1}{2}(b+c-e-f)$ The atomic p_z orbitals are designated here as a, b, c, d, e, f, in order round the ring The numbers in parentheses below the structures give the numbers of equivalent structures of the type indicated, which enter with identical coefficients

and so on, do indeed occur, but their coefficients are much smaller than those of the covalent structures and than that of

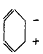


Moreover, the occurrence of resonance between

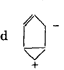
and



for example, must contribute considerably to the stability of the molecule, since this can be thought of as resulting in the formation of a three-electron bond between the carbon atoms at the extreme right of the figures The situation is somewhat analogous to that obtaining in the He_2^+ molecule-ion, which can be considered to resonate similarly between the structures He He^+ and He^+ He (Pauling 1931) In the same sense, resonance between



and



is equivalent to a one-electron bond, analogous to that in the H_2^+ molecule ion The importance of these observations lies in the fact that with cyclobutadiene such one- and three-electron bonds can be formed (as with benzene) in the triplet state, Φ_4, but not in the singlet state, $(\Phi_1 - \Phi_2)/\sqrt{2}$, because the necessary structures occur in the former, but not in the latter case This can be expressed in other words The molecular orbital treatment makes the singlet state of cyclobutadiene less stable than the triplet state, and relatively less stable than the ground state of benzene, because the resonance interactions among the ionic structures are less effective in the first case than in the last two

It is difficult to draw particular conclusions from these general considerations without actual calculation If the various ionic structures really are as important as the molecular orbital treatment makes them, then the ground state of cyclobutadiene must certainly be a triplet, if, on the other hand, they are as unimportant as the valence-bond treatment assumes, then the ground state must be a singlet Weinbaum's calculations (S Weinbaum 1933) for H_2 tend to support the latter view, since he found that the coefficients of the ionic terms were actually much smaller than those of the covalent, and that, accordingly the molecular energy given by the purely covalent valence-bond function was better by several tenths of an electron-volt than that given by the molecular orbital function (L Pauling and E B Wilson 1935) There is no assurance, however, that this same situation would be encountered in more complicated systems, such as those considered here, and, consequently, we can come to no definite conclusion It seems necessary, therefore, to leave open for the present the question as to the relative stabilities of the various spectroscopic states of cyclobutadiene, and also the parallel question as to the relative reliabilities of the valence-bond and molecular orbital methods

The author is indebted to Professor Lennard-Jones for much advice and criticism in the preparation of this paper

SUMMARY

The valence-bond and the molecular orbital treatments of cyclobutadiene disagree in two respects The first treatment predicts that the resonance energy should be about $-1\ 5$ e-volts and that the ground state should be a singlet, the second, on the other hand, predicts that the resonance energy should be zero and that the ground state should be a triplet The first of these discrepancies is found to be a result of the fact that the molecular orbital treatment neglects the resonance between the two Kekulé-like structures The second discrepancy is found to be due to the effect of ionic structures, which are not taken into account in the valence-bond treatment, but which are made important in the molecular orbital treatment The resonance among these ionic structures is of such a nature as to stabilize the triplet considerably more than the singlet state A similar discussion of benzene shows that the ionic structures contribute to the stability of the singlet ground state in essentially the same way as to the lowest triplet state of cyclobutadiene

REFERENCES

Huckel, E 1931 *Z Phys* **70**, 204
Lennard Jones, J E 1937 *Proc Roy Soc* A, **158**, 280
Lennard Jones, J E and Turkevitch, J 1937 *Proc Roy Soc* A, **158**, 297
Pauling, L 1931 *J Amer Chem Soc* **53**, 3225
Pauling, L and Wheland, G W 1933 *J Chem Phys* **1**, 362
Pauling, L and Wilson, E B Jr 1935 "Introduction to Quantum Mechanics", New York McGraw Hill Book Co
Penney, W G 1934 *Proc Roy Soc* A, **146**, 223
— 1937 *Proc Roy Soc* A, **158**, 306
Weinbaum, S 1933 *J Chem Phys* **1**, 593
Wheland, G W 1934 *J Chem Phys* **2**, 474

The electronic structure of some polyenes and aromatic molecules

VI—Phenylethylene, stilbene, tolane and the phenylmethyl radical

By W G Penney and G J Kynch

Imperial College, London, S W 7

(*Communicated by J E Lennard-Jones, F R S —
Received* 24 *November* 1937)

Introduction

Observation and theory agree that the internuclear separation of two carbon atoms linked together in a molecule may be anything between 1 54 A, a distance characteristic of the single bond, and 1 20 A, the triple bond distance Deviations from the single, double and triple bond distances are especially to be expected when neighbouring carbon-carbon links in the molecule are of different order An approximate method of calculating lengths of linkages in such molecules was given by Pauling, Brockway and Beach (1935), who also applied their theory to links involving atoms other than carbon Lennard-Jones (1937*a*), Lennard-Jones and Turkevitch (1937*b*), and Penney (1937), in papers I, II and III of the present series, have gone into the particular matter of carbon bonds more thoroughly, and have considered in detail the molecules of benzene, butadiene, hexatriene, biphenyl and napthalene, and the crystal graphite Unfortunately, extensive comparison with experiment was not possible because of the lack of measurements, but in those cases where it could be made, namely benzene and graphite, the agreement was excellent

Very recently, Robertson and Woodward (1937) have made careful X-ray analyses of crystals of stilbene and tolane, and have succeeded in determining with great accuracy the dimensions of the stilbene and tolane molecules * Their results are very interesting from a theoretical point of view because they offer the first possibility of a stringent test for the theory

---

* We wish to express our thanks to Dr Robertson and Miss Woodward for informing us of their results before publication

As we shall explain, the dimensions of the stilbene and tolane molecules may be easily predicted when accurate calculations have once been made on the phenylethylene molecule, in which the essential feature is a single bond separating a double bond and an aromatic ring Robertson and Woodward find that the length of the "single" bond in stilbene is 1 44 A, and is therefore very much shorter than the usual single bond Even shorter is the "single" bond in tolane, the length being 1 41 A The theory which we develop gives by two different methods of approximation the values 1 44 A in stilbene, and 1 41 A in tolane In addition, we find that distances within the ring are with high approximation the same as those in benzene, and that the $C{=}C$ and $C{\equiv}C$ links are practically the same as the corresponding links in ethylene and acetylene respectively These conclusions are also borne out by experiment

When constructing a theory to represent the structure of a system such as $C{=}C{-}C_6H_5$, one first assumes that most of the bonds are localized, and that the effect of these is the same as that of similar bonds in molecules where they alone are present There remain eight valencies, one on each carbon atom The net effect of these is to form a "hyper-bond", whose bonding strength is characteristic of the system, and is different from that of any classical bond structure The chief mathematical problem is thus to calculate the most stable state of an eight-electron system, the electrons being located in eight similar atomic orbitals of the carbon atoms Two methods of approximation are available, that of electron pairs and that of molecular orbitals Section I of the present paper is based on the first method, and employs the results previously obtained in paper III, section II is based on the second method, and follows closely the calculations of papers I and II

I THE APPROXIMATION OF ELECTRON PAIRS

The energy W of the most stable state of a $2n$-electron system, according to the pair method, may be obtained by solving a determinantal equation of degree $2n!/n!(n+1)!$, in which each element is a linear combination of terms in W and the exchange integrals between the electrons, taken in pairs The full equation, however, may often be simplified when the system under consideration has any symmetry Thus, the eight-electron system with which we shall deal is symmetrical about the central plane In this case, instead of beginning with an equation of order 14, one of order 10 may be substituted The ten types of canonical structure used for setting up the equation are shown in fig 1 Some of the structures, it will be seen, are labelled by two numbers, the numbers refer to the structure shown and its mirror image in the central plane

Because of the symmetry of the system about the central plane, there are only five independent exchange integrals Let Ψ_a be the accurate wave-function for the most stable state in terms of the wave-functions of the canonical structures i

$$\Psi_a = \Sigma_i a_i \psi_i, \tag{1}$$

in which the coefficients a_i are all functions of the five different exchange integrals The secular equation resulting from the elimination of the a_i, of

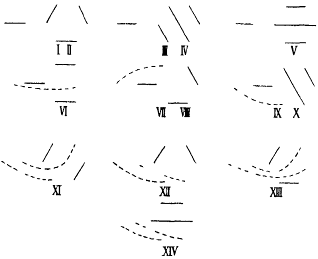


Fig 1—The fourteen canonical structures for phenylethylene

order 10, is too difficult to solve without introducing approximations As a first attempt we assume that all five exchange integrals are equal, and equal to J The equation is still of order 10, but now depends on only one parameter instead of five By an approximation of a different type, described in the following paragraphs, this equation can be replaced, without appreciable loss of accuracy, by an equation of order 6 which can be solved as closely as desired We then have the energy W as a numerical multiple of J, and can now obtain an approximate expression Ψ for the complete wave-function of (1)

As a better approximation, we take for the exchange integrals $(J+a)$, $(J+b)$, $(J+c)$, $(J+d)$ and $(J+e)$, where a, b, c, d and e are small compared with J and refer to the link A, the link B, the links C, and so on

The energy of the most stable state, accurately to the first order in a, b, c, d and e, is then calculated by first order perturbation theory

$$W = \left\{\int \Psi H \Psi dv\right\} \Big/ \left\{\int \Psi \Psi dv\right\} \tag{2}$$

The orders of the linkages as defined in paper III then follow at once, and thence x_a, x_b, x_e, the lengths of the linkages A, B, E

The approximations by which the full equation of order 10 in the single parameter J is reduced to a sixth order equation in J is as follows We assume that in the complete wave-function Ψ_a, as defined by (1), the doubly excited structures IX, XIII all have the same a_i-coefficient, and that the triply excited structure XIV can be neglected altogether There are now only six independent parameters in (1) and their elimination leads to the sixth order secular equation already mentioned

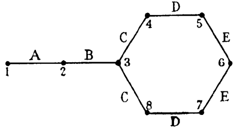


FIG 2—The lengths of the links A, B, E are denoted in the text by x_a, x_b, x_e

Approximations similar in nature to those just described, but more drastic in degree, have often been made by Pauling and his collaborators, in order to solve the secular equations for systems much more complicated than the one which we are at present considering Thus all exchange integrals in any problem are assumed equal and, for example, the equation for naphthalene, of order 16, is reduced to a cubic The stability of naphthalene and other even larger molecules as calculated from these simple equations agrees very well with the thermochemical data The principal reason for this agreement is that the particular choice of the coefficients a_i, obtained by solving the full equation in the single parameter J, minimizes the energy W as defined by (2) (For a proof of this statement see, for example, the book of Pauling and Wilson (1935), p 185) For the most stable state all the a_i's are positive Clearly, replacing each of the a_i for the set of similar excited structures by something which is very nearly their mean, causes only a small error in W The effect is to replace those terms of (1) which refer to a particular set of excited structures by a single term $\bar{a}_i \Sigma_i \psi_i$, the summation being over the set, and $\bar{a}_i$ being very close to the average value of the a_i's for the set Those rows and columns of the secular equation in J referring to the ψ_i's may now be replaced by a single row and column referring to $\Sigma_i \psi_i$ The order of the secular equation in J has been appreciably reduced

Mathematical details

The wave-functions which we use as a basis for setting up the approximate sixth order secular equation are as follows

$$\psi_1 = \mathrm{I}+\mathrm{II}, \qquad \psi_4 = \mathrm{VI},$$
$$\psi_2 = \mathrm{III}+\mathrm{IV}, \qquad \psi_5 = \mathrm{VII}+\mathrm{VIII},$$
$$\psi_3 = \mathrm{V}, \qquad \psi_6 = \mathrm{IX}+\mathrm{X}+\mathrm{XI}+\mathrm{XII}+\mathrm{XIII},$$

where, for example, I means the wave-function for the canonical structure I

The various elements of the secular equation

$$\mathrm{Det}\ \{H(k,l)-\Delta(k,l)\,W\} = 0, \quad (k=1, \quad 6,\ l=1, \quad 6)$$

are as follows

$$\{1,1\} = \tfrac{29}{4}J+\tfrac{5}{2}a-\tfrac{5}{4}b+2c+2d+2e-\tfrac{5}{2}W,$$
$$\{1,2\} = 7J+2a-b+\tfrac{5}{2}c+d+\tfrac{5}{2}e-2W,$$
$$\{1,3\} = \tfrac{7}{2}J+a-\tfrac{1}{2}b+\tfrac{1}{2}c+2d+\tfrac{1}{2}e-W,$$
$$\{1,4\} = \tfrac{5}{2}J+\tfrac{1}{2}a+\tfrac{1}{2}b+\tfrac{1}{4}c+d+\tfrac{1}{4}e-\tfrac{1}{2}W,$$
$$\{1,5\} = \tfrac{11}{2}J+\tfrac{5}{4}a+\tfrac{5}{4}b+c+d+e-\tfrac{5}{4}W,$$
$$\{1,6\} = \tfrac{89}{8}J+\tfrac{11}{4}a+\tfrac{7}{8}b+\tfrac{23}{8}c+\tfrac{7}{4}d+\tfrac{23}{8}e-\tfrac{11}{4}W,$$
$$\{2,2\} = \tfrac{17}{4}J+\tfrac{5}{2}a-\tfrac{5}{4}b+2c-d+2e-\tfrac{5}{2}W,$$
$$\{2,3\} = \tfrac{13}{4}J+\tfrac{1}{2}a-\tfrac{1}{4}b+c+d+e-\tfrac{1}{2}W,$$
$$\{2,4\} = 2J+\tfrac{1}{4}a+\tfrac{1}{4}b+\tfrac{1}{2}c+\tfrac{1}{2}d+\tfrac{1}{2}e-\tfrac{1}{4}W,$$
$$\{2,5\} = 5J+a+b+\tfrac{5}{4}c+\tfrac{1}{2}d+\tfrac{5}{4}e-W,$$
$$\{2,6\} = \tfrac{73}{8}J+\tfrac{23}{8}a+b+\tfrac{11}{4}c-\tfrac{1}{4}d+\tfrac{11}{4}e-\tfrac{23}{8}W,$$
$$\{3,3\} = \tfrac{1}{2}J+a-\tfrac{1}{2}b-c+2d-e-W,$$
$$\{3,4\} = J+\tfrac{1}{2}a+\tfrac{1}{2}b-\tfrac{1}{2}c+d-\tfrac{1}{2}e-\tfrac{1}{2}W,$$
$$\{3,5\} = \tfrac{5}{2}J+\tfrac{1}{2}a+\tfrac{1}{2}b+\tfrac{1}{4}c+d+\tfrac{1}{4}e-\tfrac{1}{2}W,$$
$$\{3,6\} = \tfrac{41}{8}J+\tfrac{7}{8}a+\tfrac{1}{2}b+c+\tfrac{7}{4}d+e-\tfrac{7}{8}W,$$
$$\{4,4\} = \tfrac{1}{2}J-\tfrac{1}{2}a+b-c+2d-e-W,$$
$$\{4,5\} = 2J-\tfrac{1}{2}a+b-c+2d+\tfrac{1}{2}e-W,$$
$$\{4,6\} = \tfrac{37}{8}J-\tfrac{1}{2}a+\tfrac{11}{8}b+\tfrac{1}{8}c+\tfrac{11}{4}d+\tfrac{7}{8}e-\tfrac{11}{8}W,$$
$$\{5,5\} = \tfrac{11}{4}J-\tfrac{5}{4}a+\tfrac{5}{2}b-\tfrac{5}{2}c+2d+2e-\tfrac{5}{2}W,$$
$$\{5,6\} = \tfrac{35}{4}J-\tfrac{7}{8}a+\tfrac{29}{8}b-\tfrac{5}{8}c+\tfrac{11}{4}d+\tfrac{31}{8}e-\tfrac{29}{8}W,$$
$$\{6,6\} = \tfrac{41}{4}J-\tfrac{11}{4}a+\tfrac{5}{2}b+\tfrac{7}{2}c-\tfrac{7}{4}d+\tfrac{35}{4}e-\tfrac{43}{4}W$$

To obtain a zero-order approximation of this equation in the special case $a = b = c = d = e = 0$, there are two obvious possibilities The first is to assume that in the complete wave-function (1), the a_i-coefficients of ψ_2, ψ_3, ψ_4 and ψ_5 are the same There results a cubic equation whose basis is

$$\psi_1, \quad (\psi_2+\psi_3+\psi_4+\psi_5), \quad \psi_6$$

The solution of this equation is $W = 3\ 3077J$, and the corresponding approximation to Ψ is

$$\left.\begin{aligned} \Psi &= \psi_1 + 0\ 3197(\psi_2+\psi_3+\psi_4+\psi_5) + 0\ 0629\psi_6, \\ \int \Psi\Psi dv &= 7\ 760 \end{aligned}\right\} \qquad (3)$$

The other possibility is that taken by Pauling and Sherman (1933), and used as a basis ψ_1, $(\psi_2+\psi_3)$ and $(\psi_4+\psi_5)$, neglecting ψ_6 The energy then works out at $3\ 3062J$, and the wave-function is

$$\left.\begin{aligned} \Psi &= \psi_1 + 0\ 4478(\psi_2+\psi_3) + 0\ 3405(\psi_4+\psi_5), \\ \int \Psi\Psi dv &= 8\ 605 \end{aligned}\right\} \qquad (4)$$

At this stage it becomes clear why calculations as elaborate as the one we are describing in this paper are necessary to find the lengths of the bonds The energy itself, because of the minimum principle, is very insensitive to the choice of an approximate wave-function, but the corresponding estimate of the contributions of the various bonds to the energy may apparently vary considerably Thus, by comparing (3) and (4), it is seen that a change of 25 % in some of the leading coefficients changes the energy by only 0 03 % By no means can one assume without close investigation that the coefficients of a, e are also insensitive to the same extent

To find a more accurate value of W, and a better approximation to Ψ, we make use of the variation principle Thus we improve on (3) by writing

$$\Psi = \psi_1 + (0\ 3197+\theta)\psi_2 + 0\ 3197(\psi_3+\psi_4+\psi_5) + 0\ 0629\psi_6,$$

and evaluate W as a power series in θ, stopping at terms in θ^3 We now choose θ to minimize W Similarly, the other coefficients may be adjusted in succession, and the process repeated until finally a wave-function is obtained in which every coefficient is accurate to within 1 %

$$\left.\begin{aligned} \Psi &= \psi_1 + 0\ 470\psi_2 + 0\ 435\psi_3 + 0\ 223\psi_4 + 0\ 363\psi_5 + 0\ 063\psi_6, \\ \int \Psi\Psi dv &= 9\ 3278, \quad W = 3\ 31901 \end{aligned}\right\} \qquad (5)$$

The value of W inclusive of first order terms in a, e may now be calculated from (2), and the resulting coefficients of a, e are accurate to the third decimal place

$$W = 3\ 319J + 0\ 878a - 0\ 037b + 0\ 730c + 0\ 901d + 0\ 847e$$

The orders of the linkages p, and the corresponding lengths x, may be calculated as described in paper III The results are summarized in the table

Link	A	B	C	D	E
p	1 919	1 309	1 577	1 634	1 616
x	1 34	1 45	1 398	1 390	1 393

Rougher approximations, etc

We have calculated the lengths x by using the approximate functions (3) and (4) With (3) the lengths x_a to x_e in Ångstroms are 1 34, 1 45, 1 400, 1 375, 1 394 With (4) the lengths are 1 34, 1 45, 1 400, 1 383, 1 393 Both sets of results are extremely good, and are far better than would ever be expected from a superficial comparison of (3), (4), (5) However, errors of about 0 02 A are present One would therefore anticipate that approximations similar to (3) or (4) for systems more complicated than an eight-electron system may easily lead to serious errors Thus, the lengths of the linkages in naphthalene as calculated in paper III may be in error by as much as 0 05 A Since the total variation found in the lengths was only 0 04 A, little reliance can be placed on the detailed values More accurate calculations are clearly needed for this molecule

Another interesting calculation is to find the lengths of the linkages by the method proposed by Pauling, Brockway and Beach (1935) Using the three functions (3), (4) and (5), the A link in each case is 1 34 A, and the ring distances are very close to the benzene ring distance The values for the B link are 1 48, 1 48 and 1 49, respectively These do not compare well with the experimental value 1 44 in stilbene

II—The orbital method

The orbital method provides a means of determining an approximate expression for $\mathscr{F}$, the energy of the "mobile" electrons together with the variation of the localized bonds, as a function of the internuclear distances By minimizing $\mathscr{F}$, the equilibrium configuration may be found, and hence the lengths of the linkages This line of approach has been developed by Lennard-Jones and Turkevitch in papers I and II of this series, where the method and nomenclature which we shall use is given in detail Our treatment is essentially similar, but is slightly more elaborate, as $\mathcal{F}$ is now a function of five variables x_a, x_b, x_c, x_d, x_e instead of only one or two

As in paper I, we assume that the molecular orbitals may be written as the sum of the wave-functions ψ_i of the individual atoms, with constant coefficients c_i,

$$\Psi = \Sigma_i c_i \psi_i, \quad (i = 1, 2, \quad 8)$$

From this equation in the usual way a secular equation is obtained whose elements are functions of the resonance and Coulomb integrals β_a, α_a, β_b, α_b, β_c, α_c, β_d, α_d, β_e, α_e, which depend respectively upon the five unknown distances x_a, x_b, x_c, x_d and x_e (see fig 2) This secular equation is of the eighth order and has the form

$$\left| \begin{matrix} E_0-E+\alpha_a, & \beta_a, & 0, & 0, & 0, & 0, & 0, & 0 \\ \beta_a, & E_0-E+\alpha_a+\alpha_b, & \beta_b, & 0, & 0, & 0, & 0, & 0 \\ 0, & \beta_b, & E_0-E+\alpha_b+2\alpha_c, & \beta_c, & 0, & 0, & 0, & \beta_c \\ 0, & 0, & \beta_c, & E_0-E+\alpha_c+\alpha_d & \beta_d, & 0, & 0, & 0 \\ 0, & 0, & 0, & \beta_d & E_0-E+\alpha_d+\alpha_e & \beta_e, & 0, & 0 \\ 0, & 0, & 0, & 0, & \beta_e, & E_0-E+2\alpha_e, & \beta_d, & 0 \\ 0, & 0, & 0, & 0, & 0, & \beta_d, & E_0-E+\alpha_e+\alpha_d, & \beta_c \\ 0, & 0, & \beta_c, & 0, & 0, & 0, & \beta_c, & E_0-E+\alpha_d+\alpha_c \end{matrix} \right.$$

To simplify this equation we make use of the symmetry of the system about a central plane Thus the introduction of zero-order wave-functions $(\psi_4 \pm \psi_8)$, $(\psi_5 \pm \psi_7)$ instead of ψ_4, ψ_5, ψ_7 and ψ_8 factorizes the equation into a sixth order equation and a quadratic equation, the quadratic being

$$\begin{vmatrix} E_0-E+\alpha_c+\alpha_d & \beta_d \\ \beta_d & E_0-E+\alpha_d+\alpha_e \end{vmatrix} = 0$$

To solve the sixth order equation, we assume that the diagonal elements of the determinant are identical and equal to ϵ This approximation is justified because the α's are small compared with the β's, and the β's are later calibrated by comparison with experiment as if the α's actually were zero The expansion then gives a cubic equation in ϵ^2

$$\begin{aligned} &\epsilon^6-\epsilon^4(\beta_a^2+\beta_b^2+2\beta_c^2+\beta_d^2+2\beta_e^2) \\ &\quad+\epsilon^2(2\beta_a^2\beta_c^2+\beta_a^2\beta_d^2+2\beta_a^2\beta_e^2+\beta_b^2\beta_d^2+2\beta_b^2\beta_e^2+4\beta_c^2\beta_e^2)-4\beta_a^2\beta_c^2\beta_e^2=0 \end{aligned}$$

The roots of this equation are

$$\epsilon_r^2 = \tfrac{1}{3}(\beta_a^2+\beta_b^2+2\beta_c^2+\beta_d^2+2\beta_e^2)+A\cos\tfrac{1}{3}(\theta+2\pi r), \quad (r=1,2,3),$$

where

$$A = \tfrac{2}{3}B^{\frac{1}{2}},$$

$$\cos\theta = \frac{C}{2B^{\frac{3}{2}}}$$

$$B = (\beta_a^2+\beta_b^2+2\beta_c^2+\beta_d^2+2\beta_e^2)^2-3(2\beta_a^2\beta_c^2+\beta_a^2\beta_d^2+2\beta_a^2\beta_e^2+\beta_b^2\beta_d^2+2\beta_b^2\beta_e^2+4\beta_c^2\beta_e^2),$$

$$\begin{aligned}C = {} & 108\beta_a^2\beta_c^2\beta_e^2+2(\beta_a^2+\beta_b^2+2\beta_c^2+\beta_d^2+2\beta_e^2)^3 \\ & -9(\beta_a^2+\beta_b^2+2\beta_c^2+\beta_d^2+2\beta_e^2)(2\beta_a^2\beta_c^2+\beta_a^2\beta_d^2+2\beta_a^2\beta_e^2+\beta_b^2\beta_d^2+2\beta_b^2\beta_e^2+4\beta_c^2\beta_e^2)\end{aligned}$$

The roots of the quadratic, with the same assumptions, are

$$\epsilon_0^2 = \beta_d^2$$

Let ϵ_r be the positive square root of ϵ_r^2 and so on, then ϵ_0, ϵ_1, ϵ_2, ϵ_3 are proportional to the various β's and these are negative Thus the four most stable orbitals have energies

$$E_r = E_0+2\alpha-\epsilon_r, \quad (r=0,1,2,3),$$

where $\epsilon_0 = \beta_d$, and ϵ_1, ϵ_2, ϵ_3 are the above roots Each of these orbitals holds two electrons, and therefore the four orbitals together will absorb all the available electrons

Physically we imagine these electrons to be placed in a framework of carbon atoms joined by single bonds and at a distance apart corresponding to a single bond After the electrons have been added, the framework has contracted so that these distances are now x_a, x_b, x_c, x_d, x_e Consequently to obtain $\mathscr{F}$ we must add to the energy of the electrons in the orbitals the "energy of compression" of the C—C distances from the single bond distance s to the equilibrium distances This is assumed to be $k_s(x-s)^2$, where k_s is *one-half* the force constant for a single bond C—C distance, determined from vibration spectra The total energy is therefore

$$\begin{aligned}\mathscr{F} = {} & 2\sum_{r=0}^{3}E_r+k_s(x_a-s)^2+k_s(x_b-s)^2+2k_s(x_c-s)^2 \\ & +2k_s(x_d-s)^2+2k_s(x_e-s)^2\end{aligned}$$

The equilibrium distances are now found by making $\mathscr{F}$ a minimum with regard to the x_i's This is equivalent to solving the equations

$$\frac{d\mathscr{F}}{dx_i} = \frac{\partial\mathscr{F}}{\partial x_i}+\frac{\partial\mathscr{F}}{\partial\beta_i}\frac{d\beta_i}{dx_i} = 0 \tag{6}$$

Each of these contains one of the distances x_i explicitly, and the integrals β_a, β_e

We assume initially that all the distances are equal, and hence $\beta_a = \beta_b = \beta_c$, etc (only the ratio of these integrals is required), and calculate the equilibrium distances using equations (6) Using these distances, we determine the resonance integrals as in paper I, and again calculate the equilibrium distances These in the second approximation were found to agree with those of the first approximation, except in the case of x_b Further approximations to x_b oscillated between 1 39 A and 1 48 A without converging However, when the mean of these, 1 44 A, was used in the calculations as a *first* attempt at x_b, the same value was obtained in the next approximation, showing that the process was now convergent This example leads us to suspect that, in certain cases, successive approximations for the lengths of the linkages will only be convergent within quite a narrow range on either side of the true value

The equilibrium distances were found to be

$$\begin{aligned} x_a &= 1\ 34\,\text{A}, & x_d &= 1\ 377\,\text{A}, \\ x_b &= 1\ 44\,\text{A}, & x_e &= 1\ 374\,\text{A} \\ x_c &= 1\ 374\,\text{A}, \end{aligned}$$

The two distances x_a and x_b differ only by 0 01 A from the values obtained by the pair method, and the other distances by 0 02 A Since the C—C distance obtained by Lennard-Jones and Turkevitch in paper II for the benzene ring was 1 37 A, both methods agree that the distances in the ring are unaltered

The small differences between the results of the pair method and of the orbital method has been suggested by Lennard-Jones in paper I as possibly due to an error in the double bond C—C distance To some extent this is confirmed, as Pauling and Brockway (1937), from new electron diffraction measurements, find 1 34 A for the double-bond distance, instead of the older value of 1 33 A This difference of 0 01 A covers the discrepancies between the lengths obtained from experiment and from the two theoretical methods

Stilbene and tolane

We have now shown by two independent methods that the lengths of the "single" bond in phenylethylene separating the double bond and the aromatic ring is 1 44 or 1 45 A, that the double bond is slightly greater (perhaps 0 01 A) than the double bond in ethylene, and that the ring distances are the same as those in benzene Thus, the substitution of a ring for a hydrogen atom in ethylene leaves the double bond practically unaffected. The lengths of the links given above must therefore apply without modification to the stilbene molecule $C_6H_5 \cdot CH{=}CH \cdot C_6H_5$. As mentioned in the introduction, Robertson and Woodward (1937) from their measurements find exactly the values which we have obtained.

Another molecule to which our results may be applied is tolane $C_6H_5 \cdot C{\equiv}C \cdot C_6H_5$. The central link again must be practically unaffected and therefore of length 1·20 A. To obtain the length of the "single" links we must consider an effect additional to those already discussed. The length of a single C—X bond depends on whether the carbon atom also participates in three other single bonds, a double and a single bond, or a triple bond. Thus in the first case there is sp^3 hybridization, in the second sp^2 and in the third sp hybridization. Because the s charge distribution is more compact than the p, or for some other cause not understood, the C—H link in ≡C—H is about 0·04 A shorter than it is in CH_4 or =C—H. A fuller discussion of the effect is to be found in an article by Penney (1935). Accordingly, we anticipate that the "single" links in tolane are 0·04 A shorter than the corresponding links in stilbene, i.e. of length 1·40 A or 1·41 A. Robertson and Woodward, in the following paper (p. 436), find 1·405 A.

Phenylmethyl radical

As a matter of interest, we have calculated by both methods of approximation the internuclear distances in the phenylmethyl radical $H_2C \cdot C_6H_5$. The A link has now disappeared, and for the links B, C, D and E, we find

$$x_b = 1\cdot 38\,\text{A}, \qquad x_d = 1\cdot 38\,\text{A},$$
$$x_c = 1\cdot 41\,\text{A}, \qquad x_e = 1\cdot 39\,\text{A},$$

as a mean of the two methods.

SUMMARY

Calculations are made to find the internuclear distances in molecules containing the system $C{=}C{-}C_6H_5$. The length obtained for the double bond is the same as that in ethylene, the ring distances are the same as those in benzene, but the "single" bond has the peculiarly short length 1·44 A. Similarly, in the system $C{\equiv}C{-}C_6H_5$, the triple bond has a normal length 1·20 A, the ring distances are the same as in benzene, but the "single" bond has a length 1·41 A. Measurements by Robertson and Woodward on stilbene and tolane agree exactly with these results.

REFERENCES

Lennard Jones 1937 *Proc Roy Soc* A, **158**, 280
Lennard Jones and Turkevitch 1937 *Proc Roy Soc* A, **158**, 297
Pauling, Brockway and Beach 1935 *J Amer Chem Soc* **57**, 2705
Pauling and Brockway 1937 *J Amer Chem Soc* **59**, 1223
Pauling and Sherman 1933 *J Chem Phys* **1**, 679
Pauling, L and Wilson, E B, Jr 1935 "Introduction to Quantum Mechanics", New York McGraw Hill Book Co
Penney 1935 *Trans Faraday Soc* **31**, 734
— 1937 *Proc Roy Soc* A, **158**, 306
Robertson and Woodward 1937 *Proc Roy Soc* A, **162**, 568
Robertson and Woodward in the following paper, p. 435

The rotational energy-levels of a diatomic molecule in a tetrahedral field

By H M Cundy, *Trinity College, Cambridge*

(*Communicated by J E Lennard-Jones, F R S — Received* 24 *November* 1937)

I—Introduction

The rotation of molecules in fields with octahedral symmetry, which is an approximation to the fields in certain crystals, has been investigated by Devonshire (1936), in a paper in which he shows the types of the wave-functions for the general field of octahedral symmetry, and then obtains energy-levels for a particular field of this type

In this paper we shall use Devonshire's method to determine the form of the wave functions and energy-levels for the similar case when the potential field has not octahedral, but only tetrahedral symmetry, or complete (A_1) symmetry of the four-point group T_d We shall again take a special form for the potential field, and show how the rotational levels split up as the field is increased, and also how the field modifies the wave-functions corresponding to the various levels

This analysis might be applied to the mathematically similar, though physically very dissimilar, case of the electronic energy-levels of tetrahedral molecules such as methane, CH_4, which are intractable by the Heitler and London method except in the ground state (Slater 1931)

II—ROTATIONAL WAVE-FUNCTIONS

We consider the rotation of a diatomic molecule in a field of potential energy V possessing tetrahedral symmetry We shall call the tetrahedron of points on a unit sphere at which V has its minimum value the *fundamental* tetrahedron, and the tetrahedron of points antipodal to these the *conjugate* tetrahedron Referred to the axes through the mid-points of opposite sides of the fundamental tetrahedron, the simplest spherical harmonic to take for the potential energy is $P_3^2(\cos\theta)\sin 2\phi$, which is the only surface harmonic of degree less than four having the required symmetry (Bethe 1929) Referred to an axis of symmetry through a vertex this becomes

$$10/\sqrt{3}\;\; P_3^0(\cos\theta)+\sqrt{2}/(3\sqrt{3})\;\; P_3^3(\cos\theta)\cos 3\phi$$

Taking the first set of axes and choosing $V = -K\;\; P_3^2(\cos\theta)\sin 2\phi$, we find that V has a minimum value of $-10K/\sqrt{3}$ at each vertex of the fundamental tetrahedron, a maximum value of $10K/\sqrt{3}$ at each vertex of the conjugate tetrahedron, and a col with value zero at the ends of the co-ordinate axes

Since V can also be written as

$$-30K\sin^2\theta\cos\theta\sin\phi\cos\phi$$

or $-30Kxyz/r^3$, it is clear that $V = 0$ on the three great circles $x = 0$, $y = 0$, $z = 0$ Fig 1 shows the orthogonal projection of the equipotential lines on the plane $z = 0$, and fig 2 the conical projection on the face $x+y+z = r/\sqrt{3}$ of the conjugate tetrahedron The projection on a face of the fundamental tetrahedron will be the same with the sign of K changed

According to the general principles of the Theory of Groups, this special assumption of a particular form for the potential energy is sufficient to enable us to find the form of the wave-functions for any potential-energy function with the same symmetry For the wave-functions are solutions of Schrodinger's equation for the diatomic molecule of moment of inertia I

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}+\frac{8\pi^2 I}{h^2}(E-V)\psi = 0, \tag{1}$$

where V is a function of θ and ϕ with tetrahedral symmetry Since this equation is invariant under the operations of the four-point group T_d, the solutions must belong to representations in this group The following table shows these representations in the notation due to Mulliken (1933)

TABLE I—REPRESENTATIONS OF THE T_d GROUP

	Characters				
Type	E	$3C_2$	$6\iota C_4$	$6\iota C_2$	$8C_3$
A_1	1	1	1	1	1
A_2	1	1	−1	−1	1
E	2	2	0	0	−1
T_1	3	−1	1	−1	0
T_2	3	−1	−1	1	0

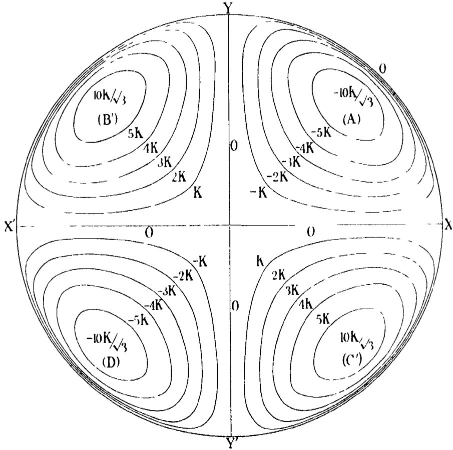


FIG 1—Orthogonal projection of the equipotentials on the plane $z = 0$

The rest of the work is brought into line with that of Devonshire (1936) by noticing that there is the following connexion between the representations of the T_d and its associated O_h group (Mulliken 1933, Table IV)

TABLE II—CORRESPONDENCE BETWEEN REPRESENTATIONS OF T_d AND O_h GROUPS

O_h	$A_{1g} A_{1u} A_{2g} A_{2u}$	$E_g E_u$	$T_{1g} T_{1u} T_{2g} T_{2u}$
T_d	$A_1\ A_2\ A_2\ A_1$	$E\ E$	$T_1\ T_2\ T_2\ T_1$

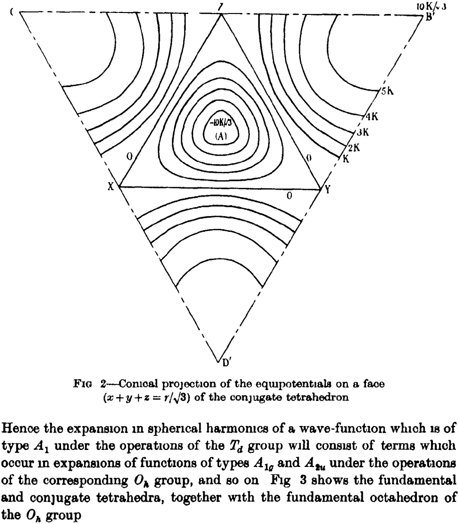


FIG 2—Conical projection of the equipotentials on a face ($x+y+z = r/\sqrt{3}$) of the conjugate tetrahedron

Hence the expansion in spherical harmonics of a wave-function which is of type A_1 under the operations of the T_d group will consist of terms which occur in expansions of functions of types A_{1g} and A_{2u} under the operations of the corresponding O_h group, and so on Fig 3 shows the fundamental and conjugate tetrahedra, together with the fundamental octahedron of the O_h group

As in the octahedral case, the solutions for $V = 0$ are the spherical harmonics, all the harmonics of degree l corresponding to the energy $E = l(l+1)$, this degeneracy is partially removed by the tetrahedral field, and the solutions split up into the following types

TABLE III—TYPES OF ROTATIONAL WAVE-FUNCTIONS IN A TETRAHEDRAL FIELD

l	
0	A_1
1	T_2
2	$E+T_2$
3	$A_1+T_1+T_2$
4	$A_1+E+T_1+T_2$
5	$E+T_1+2T_2$
6	$A_1+A_2+E+T_1+T_2$
7	$A_1+E+2T_1+2T_2$
8	$A_1+2E+2T_1+2T_2$
9	$A_1+A_2+E+2T_1+3T_2$
10	$A_1+A_2+2E+2T_1+3T_2$
11	$A_1+2E+3T_1+3T_2$
12	$2A_1+A_2+2E+3T_1+3T_2$
13	$A_1+A_2+2E+3T_1+4T_2$

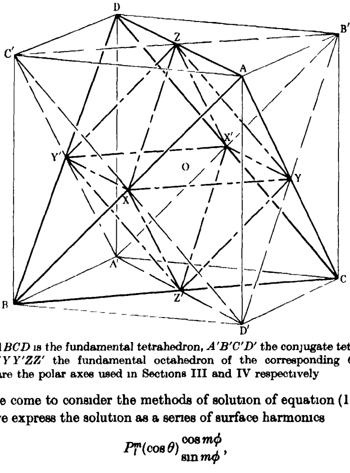


FIG 3—$ABCD$ is the fundamental tetrahedron, $A'B'C'D'$ the conjugate tetrahedron, and $XX'YY'ZZ'$ the fundamental octahedron of the corresponding O_h group OZ, OA are the polar axes used in Sections III and IV respectively

When we come to consider the methods of solution of equation (1) we find that if we express the solution as a series of surface harmonics

$$P_l^m(\cos\theta)\,{\cos m\phi \atop \sin m\phi},$$

it will contain either terms with m odd or terms with m even, but not both, this arises from the fact that V contains only terms with m even Hence the solutions divide off into two types at once, those with m odd cannot be further subdivided by elementary considerations of group theory, since, however, all the energy-levels can be found from the solutions with even m, the wave-functions with m odd do not lead to anything new, and we shall not consider them further The series with m even can be partially separated by using the subgroup D_{4h} (compare Table IV, Devonshire 1936) into the following independent types

TABLE IV—TYPES OF SOLUTION OF EQUATION (1) AND THEIR GROUP REPRESENTATIONS

T_d	D_{4h}	Series
A_1 or E	A_{1g} and B_{1u}	$a_0^0 + a_2^0 P_2^0 + a_4^0 P_4^0 + a_4^4 P_4^4 \cos 4\phi +$ $+ b_3^2 P_3^2 \sin 2\phi + b_5^2 P_5^2 \sin 2\phi +$
A_2 or E	A_{1u} and B_{1g}	$a_2^2 P_2^2 \cos 2\phi + a_4^2 P_4^2 \cos 2\phi + a_6^2 P_6^2 \cos 2\phi + a_6^6 P_6^6 \cos 6\phi +$ $+ b_5^4 P_5^4 \sin 4\phi + b_7^4 P_7^4 \sin 4\phi +$
T_1	A_{2g} and B_{2u}	$a_3^2 P_3^2 \cos 2\phi + a_5^2 P_5^2 \cos 2\phi +$ $+ b_4^4 P_4^4 \sin 4\phi + b_6^4 P_6^4 \sin 4\phi +$
T_2	A_{2u} and B_{2g}	$a_1^0 P_1^0 + a_3^0 P_3^0 + a_5^0 P_5^0 + a_5^4 P_5^4 \cos 4\phi +$ $+ b_2^2 P_2^2 \sin 2\phi + b_4^2 P_4^2 \sin 2\phi +$
T_1 or T_2	E_u and E_g	All terms with m odd

In the series which belong to each of two types, some of the terms can be immediately separated, e g in the first series the term a_0^0 is clearly A_1, and the term $a_2^0 P_2^0$ is of type E, but there will be two different combinations of P_4^0 and $P_4^4 \cos 4\phi$ which are of types A_1 and E respectively The combination of type A_1 is that combination which Devonshire has shown to be A_1 also under the O_h group, viz $P_4^0 + \frac{1}{168} P_4^4 \cos 4\phi$ The simplest way to find these correct linear combinations in any given case is to multiply terms known to be of the given type by $P_3^2 \sin 2\phi$, which is known to be of type A_1, and hence the resulting product is of the same symmetry type as the multiplicand This leads to the following series for the A_1 type

$$p_0 + p_3 P_3^2 \sin 2\phi + p_4 \{P_4^0 + \tfrac{1}{168} P_4^4 \cos 4\phi\}$$

$$+ p_6 \{P_6^0 - \tfrac{1}{360} P_6^4 \cos 4\phi\} + p_7 \{P_7^2 \sin 2\phi + \tfrac{1}{1560} P_7^6 \sin 6\phi\} +$$

III—ENERGY-LEVELS OF PARTICULAR EQUATION

We now consider equation (1) with the particular form of the potential energy $V = -K\ P_3^2(\cos\theta)\sin 2\phi$, and we set out to find the eigenvalues of the energy E as functions of the parameter K There are three methods open to us, corresponding to the three methods of Devonshire's paper, viz the general solution in the form of an infinite determinant, the asymptotic solution for large (positive or negative) values of K, and a solution by successive approximation for small values of K We are thus able to approximate closely to the lower energy-levels over quite large ranges of K Changing the sign of K merely amounts to interchanging the fundamental and the conjugate tetrahedra, and the whole diagram of energy-levels plotted as functions of K will therefore be symmetrical about $K = 0$ We proceed at first with the general method of solution

We require to know the expansion of $P_3^2(\cos\theta)\sin 2\phi\ P_l^m(\cos\theta)\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$ in surface harmonics Since we are concerned only with even values of m, the anomalous case of $m = 1$ need not be considered, and we require only the expansions of $P_3^2\ P_l^m$ in terms of P_s^{m+2} and P_s^{m-2} These are obtained by methods similar in all respects to those used by Devonshire, and therefore need not be set out here in full, the results obtained are as follows

$$P_3^2\ P_l^m = \frac{15(l-m+1)}{(2l+1)(2l+3)(2l+5)}\ P_{l+3}^{m+2} - \frac{15(l-3m-2)}{(2l-1)(2l+1)(2l+5)}\ P_{l+1}^{m+2}$$
$$-\frac{15(l+3m+3)}{(2l-3)(2l+1)(2l+3)}\ P_{l-1}^{m+2} + \frac{15(l+m)}{(2l-3)(2l-1)(2l+1)}\ P_{l-3}^{m+2}, \quad (2)$$

$$P_3^2\ P_l^m = \frac{15(l-m+5)(l-m+4)(l-m+3)(l-m+2)(l-m+1)}{(2l+1)(2l+3)(2l+5)}\ P_{l+3}^{m-2}$$
$$-\frac{15(l-m+3)(l-m+2)(l-m+1)(l+m)(l+3m-2)}{(2l-1)(2l+1)(2l+5)}\ P_{l+1}^{m-2}$$
$$-\frac{15(l-m+1)(l+m)(l+m-1)(l+m-2)(l-3m+3)}{(2l-3)(2l+1)(2l+3)}\ P_{l-1}^{m-2}$$
$$+\frac{15(l+m)(l+m-1)(l+m-2)(l+m-3)(l+m-4)}{(2l-3)(2l-1)(2l+1)}\ P_{l-3}^{m-2} \quad (3)$$

We now write equation (1) in the form

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} + \{W + k\ P_3^2(\cos\theta)\sin 2\phi\}\psi = 0, \quad (4)$$

where $W = 8\pi^2 IE/h^2$, and $k = 8\pi^2 IK/h^2$

Then since the solutions split up as in Table IV, we take any one of the series in that table, substitute it for ψ in the above equation, express every term as a sum of surface harmonics, and equate to zero the coefficient of each separate harmonic This leads us to a set of equations between the coefficients a and b Eliminating the a's and b's from these sets of equations, we obtain an infinite determinantal equation for W corresponding to each different type of solution in Table IV We then find an approximate solution of this determinant by taking only its first ten rows and columns, and solving it by means of Mr Mallock's electrical calculating machine *

IV—ASYMPTOTIC SOLUTION FOR K LARGE

We set out to find a solution in descending powers of $k^{\frac{1}{2}}$ for W and ψ This will converge for sufficiently large values of k From the physical standpoint it is easy to see that it corresponds to oscillation of the molecule in one of the four potential hollows at the vertices of the fundamental tetrahedron This state of affairs can be expected to continue only if W is less than its value at the cols separating the hollows, i e if W is negative This gives a necessary condition of k for the expansion to be valid

We begin by writing equation (4) expanded about an axis of symmetry through a vertex of the fundamental tetrahedron (OA in fig 3) It is

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}+\left\{W+\frac{10k}{\sqrt{3}}\left(P_3^0+\frac{\sqrt{2}}{30}P_3^3\cos 3\phi\right)\right\}\psi=0 \qquad (5)$$

Write $z=(10\sqrt{3}\,k)^{\frac{1}{2}}\sin^2\theta$, and we obtain

$$\begin{aligned}(10\sqrt{3}\,k)^{\frac{1}{2}}&\left(4z\frac{\partial^2\psi}{\partial z^2}+4\frac{\partial\psi}{\partial z}+\frac{1}{z}\frac{\partial^2\psi}{\partial\phi^2}-z\psi\right)\\&-\left(4z^2\frac{\partial^2\psi}{\partial z^2}+6z\frac{\partial\psi}{\partial z}-\tfrac{3}{8}z^2\psi\right)\\&+\frac{(5\sqrt{3}\,k/2)^{\frac{1}{4}}}{3}z^{\frac{3}{2}}\cos 3\phi\;\;\psi+(W+10k/\sqrt{3})\,\psi+O(k^{-\frac{1}{2}})=0\end{aligned} \qquad (6)$$

Now if we put $\psi=e^{-z/2}\;\;v$, we find

$$\begin{aligned}(10\sqrt{3}\,k)^{\frac{1}{2}}&\left\{4z\frac{\partial^2 v}{\partial z^2}+4(1-z)\frac{\partial v}{\partial z}+\frac{1}{z}\frac{\partial^2 v}{\partial\phi^2}-2v\right\}\\&-\left\{4z^2\frac{\partial^2 v}{\partial z^2}+2z(3-2z)\frac{\partial v}{\partial z}+(\tfrac{5}{8}z^2-3z)\,v\right\}\\&+\left\{\frac{(5\sqrt{3}\,k/2)^{\frac{1}{4}}}{3}z^{\frac{3}{2}}\cos 3\phi+W+10k/\sqrt{3}+O(k^{-\frac{1}{2}})\right\}v=0\end{aligned} \qquad (7)$$

---

* This machine, which is fully described by Mallock (1932), has now been purchased by the Mathematical Faculty at Cambridge

Now let $v = f_0 + k'^{-\frac{1}{2}}\ f_{\frac{1}{2}} + k'^{-1}\ f_1 + \quad ,$

$$W = -\frac{k'}{3} + k'^{\frac{1}{2}}\ a + k'^{\frac{1}{2}}\ a' + a_0 + k'^{-\frac{1}{2}}\ a_{\frac{1}{2}} + k'^{-1}\ a_1 + \quad ,$$

where $$k' = 10\sqrt{3}\,k$$

Substituting these expansions in the differential equation, and equating coefficients of powers of k', we obtain a set of equations the first three of which are as follows

$$Q(f_0) \equiv \left\{4z\frac{\partial^2}{\partial z^2} + 4(1-z)\frac{\partial}{\partial z} + \frac{1}{z}\frac{\partial^2}{\partial \phi^2} + (a-2)\right\} f_0 = 0, \tag{8A}$$

$$Q(f_{\frac{1}{2}}) + \left(a' + \frac{1}{3\sqrt{2}}z^{\frac{3}{2}}\cos 3\phi\right) f_0 = 0, \tag{8B}$$

$$Q(f_1) - \left\{4z^2\frac{\partial^2}{\partial z^2} + 2z(3-2z)\frac{\partial}{\partial z} + (\tfrac{5}{8}z^2 - 3z) - a_0\right\} f_0$$

$$+ \left\{a' + \frac{1}{3\sqrt{2}}z^{\frac{3}{2}}\cos 3\phi\right\} f_{\frac{1}{2}} = 0 \tag{8C}$$

The first equation is identical with equation (22 A) of Devonshire's paper, and it is there shown that a solution which is single-valued and small for large values of z is expansible in the form

$$f_0 = \sum_{r\ m} L_r^m(z)\ \{b_r^m \cos m\phi + c_r^m \sin m\phi\},$$

where r, m are positive integers and $r \geqslant m$, and the following conditions are fulfilled

$$2r - m = n, \quad a = 2(n+1)$$

$L_r^m(z)$ is the associated Laguerre polynomial defined by the formula

$$L_r^m(z) = z^{\frac{1}{2}m}\frac{d^m}{dz^m}\{L_r(z)\} = z^{\frac{1}{2}m}\frac{d^m}{dz^m}\left\{e^z\frac{d^r}{dz^r}(z^r e^{-z})\right\}$$

Hence if we fix n, and substitute for a and f_0 in equation (8 B), we obtain

$$Q(f_{\frac{1}{2}}) - \sum_{s,\,p} L_s^p(z)\ \ (g_s^p \cos p\phi + h_s^p \sin p\phi) = 0, \tag{9}$$

where the coefficients g_s^p and h_s^p are linear functions of the b's and c's obtained with the use of the recurrence formulae

$$z^{\frac{3}{2}}\ L_r^m = (r-m+3)(r-m+2)(r-m+1)\,L_r^{m-3} - 3r^2(r-m+2)(r-m+1)\,L_{r-1}^{m-3}$$
$$+\,3r^2(r-1)^2(r-m+1)\,L_{r-2}^{m-3} - r^2(r-1)^2(r-2)^2\,L_{r-3}^{m-3},$$

$$z^{\frac{3}{2}}\ L_r^m = L_r^{m+3} - \frac{3}{r+1}\ L_{r+1}^{m+3} + \frac{3}{(r+1)(r+2)}\ L_{r+2}^{m+3}$$
$$-\,\frac{1}{(r+1)(r+2)(r+3)}\ L_{r+3}^{m+3},$$

$$z^{\frac{3}{2}}\ L_r^1 = \frac{r}{(r+1)(r+2)}\ L_{r+2}^2 - \frac{3r}{r+1}\ L_{r+1}^2 + 3r\ L_r^2 - r^2\ L_{r-1}^2,$$

$$z^{\frac{3}{2}}\ L_r^2 = -\frac{r(r-1)}{r+1}\ L_{r+1}^1 + 3r(r-1)\ L_r^1 - 3r^2(r-1)\ L_{r-1}^1$$
$$+\,r^2(r-1)^2\ L_{r-2}^1,$$

which are all easily found by differentiation and successive application of well-known relations

The particular integral of the equation (9) is at once seen to be

$$f_{\frac{1}{2}} = \sum_{s\ p}\left\{\frac{1}{2(n-2s+p)}\right\}\ L_s^p\ \ (g_s^p \cos p\phi + h_s^p \sin p\phi), \tag{10}$$

the summation being carried out over all terms for which $2s-p \neq n$, together with terms with logarithmic singularities at the origin arising from those terms for which $2s-p=n$ But since we require that $f_{\frac{1}{2}}$ shall be finite at the origin, the coefficients of all these terms must vanish This leads us to $2n+1$ linear equations between the $2n+1$ coefficients b_r^m c_r^m In this case, however, an inspection of the recurrence formulae reveals the fact that the m of all the terms arising from the term $z^{\frac{3}{2}}\cos 3\phi\ f_0$ will be of opposite parity to the m's occurring in f_0, which are all of the same parity, viz that of n Hence the only terms for which $2s-p=n$ must have a p of the same parity as n, and must therefore arise from the term $a'f_0$, the above $2n+1$ equations are therefore all equivalent to the single equation $a'=0$

We now apply the procedure again to equation (8 C), and putting in the values of a and a', and remembering that $Q(f_0)=0$, we find that it can be written as

$$Q(f_1) - \left\{2z\frac{\partial}{\partial z} - \frac{\partial^2}{\partial\phi^2} - z(2n+3) + \tfrac{5}{8}z^2 - a_0\right\} f_0 + \frac{1}{3\sqrt{2}} z^{\frac{3}{2}} \cos 3\phi\ \ f_{\frac{1}{2}} = 0 \tag{11}$$

Hence $Q(f_1)$ is given in terms of known quantities for any given f_0, and the same reasoning as before leads us to a determinantal equation for a_0 Beyond this the process becomes very laborious even in the simplest cases

The work is simplified by the observation that the successive f's must all belong to one or other of the following classes, the equation (7) and the recurrence formulae being of such a nature that all the polynomials entering into the work are confined to the class in which the initial ones are taken to lie (see Table V) The solutions are all degenerate in sets of four and eight, corresponding to ψ being small except near one out of four potential minima, physically also it is clear that the molecule will have the same energy for a given type of oscillation about any of its four equilibrium positions If we take these degenerate sets of solutions together, we can determine their characters under the operations of the tetrahedral group, and hence find the symmetry classes which combine to form these asymptotic solutions We thus find the types of energy curves which tend asymptotically to the value found by this method

TABLE V—CLASSES OF ASYMPTOTIC ROTATIONAL WAVE FUNCTIONS FOR LARGE k

Solution	T_d representation
$\Sigma g_r^{3m} \quad L_r^{3m} \cos 3m\phi$	$A_1 + T_2$
$\Sigma h_r^{3m} \quad L_r^{3m} \sin 3m\phi$	$A_2 + T_1$
$\Sigma g_r^{3m\pm 1} \quad L_r^{3m\pm 1} \cos(3m \pm 1)\phi$	$E + T_1 + T_2$
$\Sigma h_r^{3m\pm 1} \quad L_r^{3m\pm 1} \sin(3m \pm 1)\phi$	

It is not possible to give a general formula for the value of a_0, but there is appended a table which gives the first three terms in the asymptotic expansion of W for low values of n The expansions cannot even be approximately valid unless k is large enough for W to be negative

TABLE VI—ASYMPTOTIC EXPANSION OF W FOR LARGE VALUES OF k

n	W	Symmetry class
0	$-10k/\sqrt{3} + 2(10\sqrt{3}k)^{\frac{1}{2}} - 1\frac{7}{9}$	$A_1 + T_2$
1	$-10k/\sqrt{3} + 4(10\sqrt{3}k)^{\frac{1}{2}} - 4\frac{4}{9}$	$E + T_1 + T_2$
2	$-10k/\sqrt{3} + 6(10\sqrt{3}k)^{\frac{1}{2}} - 7\frac{7}{9}$	$E + T_1 + T_2$
2	$-10k/\sqrt{3} + 6(10\sqrt{3}k)^{\frac{1}{2}} - 11\frac{1}{9}$	$A_1 + T_2$

Since the whole energy diagram is symmetrical about $k = 0$, the same asymptotic expansions will hold for k large and negative if we write $-k$ for k

V—ASYMPTOTIC SOLUTION FOR k SMALL

We now consider the case where the field is small enough to be treated as a small perturbation, and the molecule is approximating to the free rotator with energy levels $W = l(l+1)$

We take the equation in the form

$$(\nabla^2 + W - U)\psi = 0,$$

where
$$U = -k \quad P_3^2(\cos\theta)\sin 2\phi$$

and
$$\nabla^2 \equiv \frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2},$$

and write $W = l(l+1) + R$, hence obtaining

$$\{\nabla^2 + l(l+1)\}\psi = (-R+U)\psi \tag{12}$$

We now solve this by Picard's method of successive approximation, first of all neglecting the right hand side, and obtaining

$$\psi_1 = \sum_m a_m \quad P_l^m \cos m\phi + \sum_m b_m \quad P_l^m \sin m\phi, \tag{13}$$

where the a's and b's are arbitrary constants, next we substitute ψ_1 in the right-hand side of equation (12) and add a further term

$$\Sigma\alpha_1^m \quad P_l^m \cos m\phi + \Sigma\beta_1^m \quad P_l^m \sin m\phi,$$

so that all the surface harmonics of degree l are cancelled, and the second approximation ψ_2 is then finite at the pole It is given by a sum of harmonics, remembering that

$$\{\nabla^2 + l(l+1)\}\psi = \sum_{k,m} p_k^m \quad P_k^m \cos m\phi + \sum_{k\ m} q_k^m \quad P_k^m \sin m\phi$$

has particular integral

$$\psi = \sum_{k\ m} \frac{P_k^m}{l(l+1) - k(k+1)}(p_k^m \cos m\phi + q_k^m \sin m\phi)$$

Hence we obtain in succession the approximations $\psi_1, \psi_2, \psi_3, \quad \psi_n$ Adding together the equations satisfied by successive approximations, we find

$$\{\nabla^2 + l(l+1) + R - U\}\sum_{s=1}^{n}\psi_s$$
$$= -(-R+U)\psi_s + \sum_m P_l^m\left\{\cos m\phi\sum_{s=1}^{n-1}\alpha_s^m + \sin m\phi\sum_{s=1}^{n-1}\beta_s^m\right\} \tag{14}$$

Now $\sum_{1}^{\infty}\psi_s$ will satisfy equation (12) if it is convergent, and from this last equation a necessary condition for this is that

$$\sum_{s=1}^{\infty}\alpha_s^m = 0, \quad \sum_{s=1}^{\infty}\beta_s^m = 0, \quad \text{for all } m$$

These are $2l+1$ linear equations in the a's and b's, and also containing R (not necessarily linearly), hence the eigenvalues of R are found from the determinantal equation which is the eliminant of the a's and b's As before, the types of solution fall into the classes given in Table IV, which simplifies the work considerably We append a table giving the first two terms in W (i e the first term in R) which gives the form of the parabola in which the energy-levels plotted against k cross the axis $k = 0$

TABLE VII—ASYMPTOTIC EXPANSION OF W FOR SMALL VALUES OF THE PARAMETER k

l	W	Symmetry class
0	$-\frac{5}{7}k^2$	A_1
1	$2-\frac{5}{21}k^2$	T_2
2	$6-\frac{75}{77}k^2$	E
	$6+\frac{25}{231}k^2$	T_2
3	$12-\frac{25}{143}k^2$	A_1
	$12-\frac{125}{286}k^2$	T_1
	$12+\frac{425}{958}k^2$	T_2
4	$20+\frac{95}{143}k^2$	A_1
	$20-\frac{205}{1001}k^2$	E
	$20-\frac{45}{286}k^2$	T_1
	$20-\frac{835}{2165}k^2$	T_2
5	$30-\frac{105}{3094}k^2$	E
	$30-\frac{5}{110}k^2$	T_1
	$\left.\begin{array}{l}30-0\ 2137k^2\\30+0\ 3214k^2\end{array}\right\}$	$2T_2$

VI—GENERAL RESULTS

Table VIII gives the values of W calculated on the Mallock machine, using the method of Section III, for different values of the paramater k From this table the diagram of fig 4 has been plotted, showing the splitting of the energy-levels as k is increased Fig 4 also shows the asymptotic values for large k, found as in Section IV, and it will be noticed that there is a good fit between the curves found by the two methods in that part of the diagram where the asymptotic values have any meaning, i e for W negative The slight divergence of some of the curves from the asymptote as k increases is attributable to the failure of the approximation of taking only ten rows and columns of the infinite determinant This is particularly the case with the representation T_2, where the large number of terms of low order involved (see Table IV) and the necessary extraction of six roots as against three of each other representation caused the higher roots to crowd down and the machine to become unstable Hence the readings at $k = 10$ for T_2 must be treated as only very approximate and for this same reason the readings at $k = 10$ of the T_2 and E branches emanating from $W = 30$ could not be taken at all, and the top points of the T_1 and A_1 branches must be treated with caution The A_1 and E determinant (corresponding to the first wave-function of Table IV) was factorized for the A_1 terms before setting up on the machine, in this case only a fifth order determinant was needed, which greatly simplified the working of the machine

TABLE VIII—W FOR VALUES OF k

Class	$k = 0$	1	2	3 3̇	5	6 6̇	10	Asymptotic value at $k = 10$
A_1	0	−0 696	−2 608	−6 42	−12 35	−18 9	−33 05	−33 19
T_2	2	0 925	−1 58	−5 83	−12 07	−18 63	−32 65	−33 19
E	6	5 84	5 33	4 13	1 62	−1 55	−10 35	−9 535
T_2	6	6 01	5 68	4 52	1 65	−1	−12	−9 535
T_1	12	11 577	10 428	8 36	5 3	0 58	−8 5	−9 535
A_1	12	11 85	11 48	10 9	10 0	8 9	6 1	10 12
T_2	12	12 445	13 340	14 46	13 7	12 8	10 3	10 12
T_2	20	19 78	19 26	19	18 9	19 6	26 1?	—
E	20	19 65	19 38	18 8	16 62	15 30	10 5	13 45
T_1	20	20 19	20 66	20 75	20 25	17 8	12	13 45
A_1	20	20 63	22 22	24 94	28 35	30 9	32 6	—
T_2	30	29 71	29 694	29 08	?	21	16 9	13 45
T_1	30	29 96	30 22	29 8	28 9	30 5	33 2	—
E	30	29 85	29 70	30 6	26 9	25 3	?	—
T_2	30	30 285	31 102	32 93	30 8	27 3	?	—

The coefficients in the wave-functions have also been tabulated, but only for the lower levels, and for the functions with even m *

The error in the energy is much less than that in the wave-function, and can be taken as less than 1% in the region where k and W are both small, on the other hand, the machine itself has given readings divergent by as

---

* This table is too cumbrous to include here, but will be readily supplied on application to the author or to Professor Lennard Jones at the University Chemical Laboratory, Cambridge

much as 6% for k large and in the region of a near approach of two energy-levels belonging to the same representation

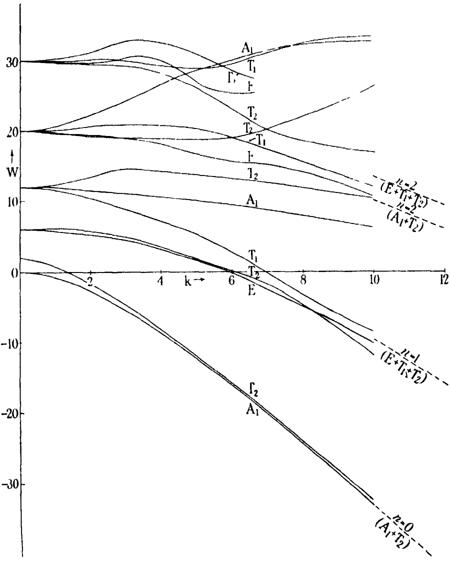


FIG 4—The asymptotic curves are shown in broken lines

I here desire to thank Professor Lennard-Jones for suggesting this problem in the first place, and for valuable help, advice and elucidation given

during the work, Dr A F Devonshire, for many hints on the application of the Group Theory to the work, Mr R R M Mallock for personal instruction in the use of his machine and the Royal Society for a grant enabling me to make use of it

SUMMARY

It is shown by general principles of the Theory of Groups how the Schrodinger equation may be solved for a diatomic molecule, when the potential energy varies with the angular co-ordinates in such a manner that it has the symmetry of the regular tetrahedron Infinite series for the wave-functions are given, and the energy-levels and corresponding wave-functions evaluated for a particularly simple potential-energy function of the given type Three methods of obtaining these levels are described, two of which are asymptotic approximations, and the third involves the use of the Mallock electric calculating machine A good agreement is obtained between the three, enabling a plot to be made of the energy levels over wide ranges of magnitude of the potential energy term

REFERENCES

Bethe 1929 *Ann Phys, Lpz*, **3**, 133
Devonshire 1936 *Proc Roy Soc* A, **153**, 601
Mallock 1932 *Proc Roy Soc* A, **140**, 457
Mulliken 1933 *Phys Rev* **43**, 278
Slater 1931 *Phys Rev* **37**, 481, **38**, 1109

X-ray analysis of the dibenzyl series

V—Tolane and the triple bond

By J Monteath Robertson, M A , D Sc and I Woodward, M A

(*Communicated by Sir William Bragg, O M , P R S —*
Received 27 *November* 1937)

The detailed structure of stilbene has recently been described (Robertson and Woodward 1937) and a theoretical discussion of the resonance problem in that type of molecule is now given by Penney and Kynch (1938)

The crystal symmetry and dimensions of a number of compounds are closely similar to those of stilbene, and it is possible to make use of this isomorphism to carry out a very direct analysis of these other structures In X-ray analysis it is generally necessary in the first place to assume some model for the structure and then adjust the parameters by a process of trial and error until a fair measure of agreement is obtained between the calculated and the observed values of the structure amplitudes Subsequent refinement of the parameter values can then be attained by the application of Fourier series methods With an isomorphous series of the present kind, however, when the structure of one member has been accurately determined, it is often possible to eliminate the first and most laborious part of the analysis for the other members, and proceed to the solution directly by the successive application of Fourier series methods

This process has been applied in the present analysis of tolane It is found that the intensities of the more outstanding X-ray reflections are similar to those of stilbene, but the agreement soon falls off when the higher orders are reached A preliminary Fourier series was accordingly set up, with the stilbene phase constants, from which the terms corresponding to the weaker and rapidly changing reflections were omitted The summation of this series gave a rough picture of the structure, from which the atomic positions could be estimated A recalculation of the structure factors from this data then enabled a more complete series to be formed, and so the analysis was completed, without the use of any molecular model or assumptions derived from the chemistry of the compound Details of the analysis, with co-ordinates, etc , are given in a later section

THE STRUCTURE OF THE TOLANE MOLECULE

As in the case of the stilbene structure, two separate tolane molecules make contributions to the asymmetric unit of the crystal, and the structures of these two molecules are deduced independently from the X-ray data Each of these molecules has a centre of symmetry, at (000) and $(00\frac{1}{2})$, and the unit cell is completed by two other molecules, at $(\frac{1}{2}\frac{1}{2}0)$ and $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, derived from the first two by reflections in the (010) plane, or, alternatively, by rotations of 180° about the *b*-axis, together with certain translations

The final results of the X-ray analysis for the two independent molecules are shown in the electron density maps of figs 1 and 2, while fig 3 shows the way in which the molecules are grouped together in the crystal These are all projections of the structure along the *b*-axis of the crystal, the direction which is found to give the best resolution of the individual atoms The diagrams should be compared with the corresponding ones for stilbene (*loc cit*) and for dibenzyl (Robertson 1935*a*)

The particular interest of the tolane structure lies in the effect of the triple bond between the central pair of carbon atoms Even a rough inspection of the diagrams, especially fig 2, is sufficient to show that a strictly linear structure has now been imposed on the central part of the molecule In dibenzyl, the angle of the central zigzag was 109–112°, in stilbene (double bond) it was 128–133°, and in tolane it can now be shown to be within 2° of 180°

To obtain the true dimensions of the molecules it is necessary first of all to calculate their orientation in the crystal, and reduce the diagrams to normal projections This is done in the next section, but we may note here that as no other favourable projections of the structure can be made, the interatomic distance in the benzene ring is first assumed to have the standard value of 1 39 A, and the orientations are calculated from this assumption The results are then checked, and the linear structure of the molecule confirmed, by calculating the more important intensities in other zones of reflections from the coordinates obtained The final results show that the difference in appearance between molecules 1 and 2 is due to a slight difference in orientation, and not to any inherent difference in shape or dimensions The single model which represents both molecules is shown in fig 4

For the carbon-carbon triple bond distance we obtain the value **1 19** A, and for the "single" bonds connecting these carbon atoms to the benzene rings we obtain the unusually small value of **1·40** A These figures are probably correct to within 0 02 A

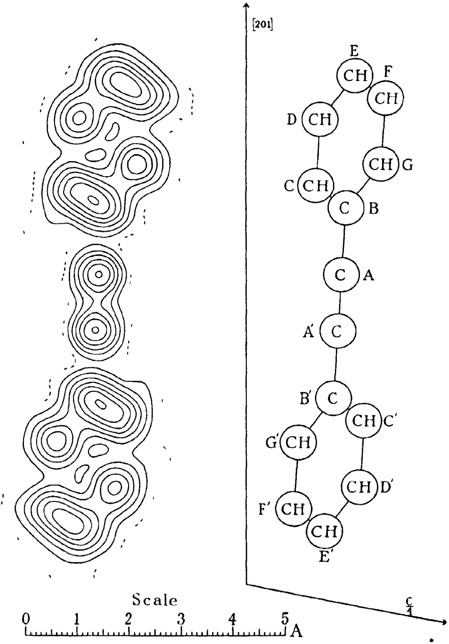


Fig 1—Tolane molecule 1 in the *b* axis projection Each contour line represents a density increment of approximately one electron per A², the one electron line being dotted

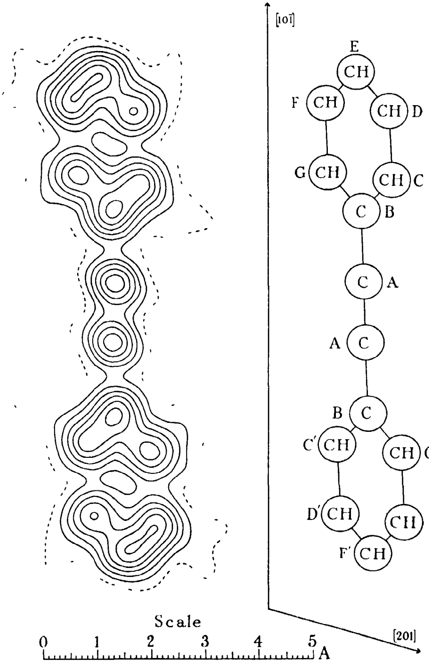


Fig 2—Tolane molecule 2 in the *b* axis projection Scale as in fig 1

There does not appear to be any other X-ray work on compounds containing a triple carbon-carbon bond, but the dimensions of the acetylene molecule have been determined very accurately by the analysis of band spectra (Mecke 1930, Herzberg, Patat and Spinks 1934), and the value obtained for the triple bond distance is 1 199 A, in good agreement with our result We might, however, have expected the tolane triple bond distance to be a little greater than the acetylene value, on account of the contribution

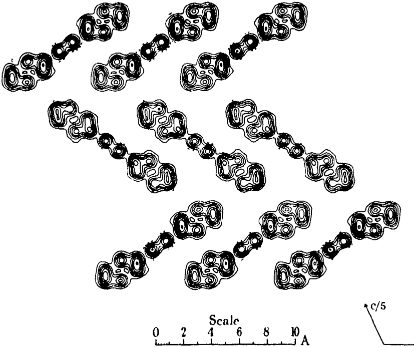


Fig 3—Group of tolane molecules in the *b* axis projection Each line approximately one electron per A^2

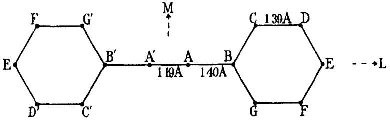


Fig 4—Dimensions of the tolane molecule

which excited structures, with a double bond between the central carbon atoms, must make to the normal state of the molecule Instead, our value is less, if anything, than that obtained for acetylene It should be noted that the distance in acetylene, obtained from the moment of inertia, gives accurately the distance between the carbon nuclei, whereas the X-ray method deals essentially with the electron distribution It does not seem unlikely that there might be a small difference in the result as given by these two methods, especially with a triple bond, due to a small inward displacement of the average electron distribution This effect, if appreciable, would tend to make the X-ray value for the distance somewhat smaller than the spectroscopic value

The "single" bond distances of 1 40 A connecting the central carbon atoms to the benzene rings, can be compared with the corresponding distances in stilbene of 1 44–1 45 A The presence of the triple bond apparently causes an extra contraction in the length of this link amounting to between 0 04 and 0 05 A, a result which it is difficult to account for in terms of the usual resonance phenomenon in aromatic and conjugated systems It would appear rather to be due to a certain increase in the *s* character of the remaining orbits of the carbon atoms, after the formation of the triple bond, the contraction being of approximately the same amount as that found for the carbon-hydrogen distance in passing from methane (1 093 A) to acetylene (1 058 A) (Ginsburg and Barker 1935, Herzberg, Patat and Spinks 1934) Further, in methyl acetylene the terminal carbon-carbon distance, adjacent to the triple bond, has recently been shown to have the abnormally small value of 1 46 A (Herzberg, Patat and Verleger 1937, Badger and Bauer 1937) The matter is further discussed by Penney and Kynch (1938)

ANALYSIS OF THE STRUCTURE

Crystal data

Tolane—$C_{14}H_{10}$, M = 178, melting point, 60° C, density, calculated, 1 134, monoclinic prismatic, $a = 12\ 75$, $b = 5\ 73$, $c = 15\ 67$ A, $\beta = 115\ 2°$ Space group, C_{2h}^5 ($P2_1/a$) Four molecules per unit cell Molecular symmetry, centre Volume of the unit cell = 1036 A^3 Absorption coefficient for X-rays, $\lambda = 1\ 54$, $\mu = 5\ 91$ per cm Total number of electrons per unit cell = F(000) = 376

Structure factors and Fourier synthesis

The measured values of the structure factors for the ($h0l$) reflections and the phase constants (or signs) employed in the final Fourier synthesis are given in Table I The absolute scale of these factors was obtained by correlation with the stilbene and dibenzyl values For the sake of brevity, the spacings of the planes and the calculated values of the structure factors, which determine the signs, are not listed The latter may be obtained from the co-ordinates of the atoms given in Table II by means of the relations

$$F(h0l) = 376 f_{CH}\, S(h0l),$$

$$14 S(h0l) = \Sigma \cos 2\pi (hx/a + lz/c),$$

the values of f_{CH} being taken from the average atomic f-curve for aromatic hydrocarbons (Robertson 1935*b*) and the summation being taken over the fourteen atoms in the asymmetric unit of the crystal It will be found that

TABLE I—VALUES OF $F(h0l)$

		h							
		0	2	4	6	8	10	12	14
	16	+5	—	—	—	—	—	—	—
	15	—	−4	—	—	—	—	—	—
	14	—	—	—	—	—	—	—	—
	13	—	—	—	—	—	—	—	—
	12	—	—	−6	—	—	—	—	—
	11	—	—	—	−7	—	—	—	—
	10	+8	−16	—	—	—	—	—	—
	9	—	+9	−14	−8	—	—	—	—
	8	−16	+18	—	—	—	—	—	—
	7	—	+11	−26	—	−9	—	—	—
	6	+15	+20	−13	—	—	—	—	—
	5	—	−3	−14	−7	−17	—	—	—
	4	−64	+9	+33	+9	—	—	—	—
	3	—	−57	+23	−34	—	—	—	—
	2	+19	+81	+31	+4	−19	—	—	—
	1	—	−51	—	−9	—	—	—	—
l	0	+376	−49	−7	—	—	—	—	—
	$\bar{1}$	—	—	−18	+24	—	−11	−7	—
	$\bar{2}$	+19	−49	+8	+13	+10	−17	+6	—
	$\bar{3}$	—	+49	+20	+4	—	−30	—	+6
	$\bar{4}$	−64	+85	—	+19	—	—	−8	—
	$\bar{5}$	—	+62	+6	−34	—	—	+8	—
	$\bar{6}$	+15	+11	+40	−7	—	+4	−24	—
	$\bar{7}$	—	—	−24	+20	—	+22	−15	—
	$\bar{8}$	−16	+17	+28	−10	—	−3	−10	—
	$\bar{9}$	—	—	+19	+37	—	—	—	—
	$\overline{10}$	+8	+12	−12	+12	−11	—	—	—
	$\overline{11}$	—	—	+20	+6	—	+10	+5	−6
	$\overline{12}$	—	−9	+9	—	—	−5	—	—
	$\overline{13}$	—	−7	+17	—	+19	—	—	—
	$\overline{14}$	—	—	—	—	—	—	—	—
	$\overline{15}$	—	+6	+6	+6	+8	−5	—	—
	$\overline{16}$	+5	−7	−7	—	—	—	—	—
	$\overline{17}$	—	+6	—	+6	—	—	—	—

the structure factors calculated in this way are in good agreement with the observed values

In addition to the structure factors listed in Table I, very weak reflections from seventeen other ($h0l$) planes were observed within the limits of the experiment (up to $\sin\theta = 0\ 87$, $\lambda = 1\ 54$), but these terms had to be omitted from the final Fourier synthesis because of doubtful sign The $(100\bar{1})$ term was also accidentally given the value − 5 instead of the measured value − 11 These omissions amount to only a very small fraction of the total values, and the Fourier series as it stands is quite reasonably convergent This series was summed at 900 points on the asymmetric unit, at intervals of 0 213 A along the a-axis and 0 261 A along the c-axis, the contour maps being prepared from the summation totals by the usual methods

In the other zones of the crystal, the reflections from the (020), (040), (024) and (032) planes are all much stronger than the corresponding stilbene values, the two latter being absent in stilbene On the other hand, the (031) and (035) reflections are weaker in tolane than in stilbene The values of the structure factors calculated from the final co-ordinates reflect these changes in every case

TABLE II—CO-ORDINATES CENTRE OF SYMMETRY AT (000) AS ORIGIN
MONOCLINIC CRYSTAL AXES

Atom	x	$2\pi x/a$	y	$2\pi y/b$	z	$2\pi z/c$
$A_1(C)$	0 63 A	17 8°	0 15 A	9 4°	0 35 A	8 0°
$B_1(C)$	2 11	59 6	0 51	31 7	1 17	26 8
$C_1(CH)$	2 47	69 7	1 68	105 5	1 99	45 7
$D_1(CH)$	3 94	111 2	2 03	127 7	2 80	64 3
$E_1(CH)$	5 06	142 8	1 21	75 8	2 80	64 2
$F_1(CH)$	4 70	132 7	0 03	2 1	1 98	45 4
$G_1(CH)$	3 23	91 2	−0 32	−20 1	1 17	26 8
$A_2(C)$	0 34	9 5	−0 06	− 3 6	7 47	171 6
$B_2(C)$	1 12	31 6	−0 19	−12 1	6 60	151 7
$C_2(CH)$	0 89	25 2	−1 30	−81 9	5 70	130 9
$D_2(CH)$	1 67	47 2	−1 44	−90 3	4 84	111 2
$E_2(CH)$	2 68	75 7	−0 46	−29 0	4 89	112 3
$F_2(CH)$	2 91	82 1	0 65	40 8	5 80	133 1
$G_2(CH)$	2 13	60 1	0 78	49 3	6 65	152 9

Orientations of the molecules Co-ordinates

The actual dimensions and orientations of the molecules were deduced from the contour maps in a very similar manner to that employed for stilbene, except that the linear structure of the tolane molecule simplifies the work The symbols have the same meaning as in the stilbene paper (Robertson and Woodward 1937)

For molecule 1 (compare figs 1 and 5) we find

$$r_{DG} = 1\ 48, \quad r_{AA} = 1\ 14, \quad r_{AB} = 1\ 35_5\,\mathrm{A},$$
$$\eta_{DG} = 89\ 5^\circ, \quad \eta_{L(A\ AB)} = 33\ 2^\circ$$

We assume that the benzene rings are regular plane hexagons with sides 1 39 A, and that the points O, A, B, E, which lie on a straight line in the projection, are actually collinear These assumptions are verified by the structure factor calculations given above Thus we take $R_{DG} = 2\ 78\,\mathrm{A}$, and the angle between L and DG to be 60° For the orientation* of molecule 1 this gives

$$\begin{array}{llll}
\chi_L = 35\ 9^\circ & \cos\chi_I = 0\ 8096 & \chi_M = 117\ 5^\circ & \cos\chi_M = -0\ 4622 \\
\psi_L = 75\ 3^\circ & \cos\psi_L = 0\ 2529 & \psi_M = 33\ 8^\circ & \cos\psi_M = 0\ 8312 \\
\omega_L = 58\ 0^\circ & \cos\omega_L = 0\ 5298 & \omega_M = 72\ 0^\circ & \cos\omega_M = 0\ 3095 \\
\chi_N = 111\ 2^\circ & \cos\chi_N = -0\ 3618 & \chi_{DG} = 89\ 7^\circ & \cos\chi_{DG} = 0\ 0046 \\
\psi_N = 119\ 7^\circ & \cos\psi_N = -0\ 4951 & \psi_{DG} = 32\ 2^\circ & \cos\psi_{DG} = 0\ 8462 \\
\omega_N = 37\ 8^\circ & \cos\omega_N = 0\ 7897 & \omega_{DG} = 57\ 8^\circ & \cos\omega_{DG} = 0\ 5329
\end{array}$$

For molecule 2 the measured quantities are

$$r_{DG} = 1\ 67, \quad r_{AA} = 1\ 20, \quad r_{AB} = 1\ 40\,\mathrm{A},$$
$$\eta_{DG} = 79\ 1^\circ, \quad \eta_{L(A\ AB)} = 34\ 1^\circ,$$

and with the same assumptions, the orientation of molecule 2 (after rotation of 180° about the a-axis) is

$$\begin{array}{llll}
\chi_L = 34\ 5^\circ & \cos\chi_L = 0\ 8243 & \chi_M = 110\ 2^\circ & \cos\chi_M = -0\ 3448 \\
\psi_L = 84\ 5^\circ & \cos\psi_L = 0\ 0965 & \psi_M = 29\ 8^\circ & \cos\psi_M = 0\ 8676 \\
\omega_L = 56\ 1^\circ & \cos\omega_L = 0\ 5582 & \omega_M = 69\ 0^\circ & \cos\omega_M = 0\ 3587 \\
\chi_N = 114\ 7^\circ & \cos\chi_N = -0\ 4490 & \chi_{DG} = 83\ 5^\circ & \cos\chi_{DG} = 0\ 1136 \\
\psi_N = 119\ 9^\circ & \cos\psi_N = -0\ 4879 & \psi_{DG} = 36\ 9^\circ & \cos\psi_{DG} = 0\ 7997 \\
\omega_N = 41\ 6^\circ & \cos\omega_N = 0\ 7482 & \omega_{DG} = 53\ 9^\circ & \cos\omega_{DG} = 0\ 5987
\end{array}$$

From these figures the true dimensions are easily obtained For molecule 1 the triple bond distance AA' is 1 18 A, and for molecule 2 it is 1 20 A For both molecules the length of AB is 1 40 A Thus the dimensions of the two molecules are found to be the same within the limits of error (about ± 0 02 A) and so we take the model shown in fig 4 to represent the final results

---

* In this and similar work (Robertson and Woodward 1937) the orientation figures are probably accurate in general to within 1°, except for movements which do not have much effect on the projections, when the error may be greater Four figure direction cosines are given to make the figures mutually consistent and to afford a convenient means of calculating other properties of the crystals, e g the magnetic susceptibilities

The co-ordinates of all the atoms in the asymmetric unit, referred to the crystal axes and the centre of molecule 1 as origin, are obtained from fig 4 and the direction cosines given above These are given in Table II, and the way in which these final co-ordinates fit the contour maps is shown in fig 5

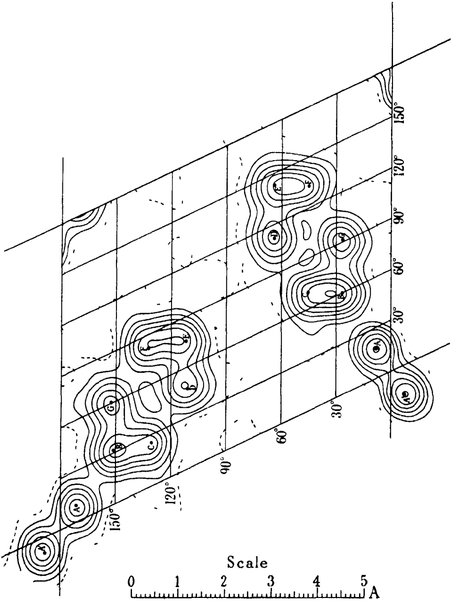


FIG 5—The co-ordinates assigned to the atoms in tolane

Intermolecular distances

The smallest intermolecular distances for the carbon atoms lie between 3 5 and 3 6 A, very similar to those of the stilbene structure The nearest approaches which have been found occur between atom *C* of molecule 1 and *G′* (centro-symmetrical to *G*) of the reflected molecule at (½½0) where the distance is **3 53** A, between *C* of molecule 1 and *E* of molecule 2, **3 54** A, between *B* of molecule 1 and *G′* of the reflected molecule at (½½0), **3·57** A, and between *B* of molecule 2 and *G′* of the reflected molecule 2 at (½½½), **3 57** A

In conclusion, we are indebted to Sir William Bragg and the Managers of the Royal Institution for laboratory facilities, to Dr J W H Oldham and Professor L Ruzicka for specimens of tolane, and to Dr A Muller and Mr H Smith for the use of a 5 kw X-ray generator

SUMMARY

The crystal dimensions of tolane are very similar to those of stilbene, and this fact has been used in making a very direct X-ray analysis of the tolane structure, without the use of any model based on the chemistry of the compound The two central atoms linked by the triple bond are found to be collinear with the end atoms of the benzene rings to which they are joined The carbon-carbon triple bond distance is **1 19** A, and for the "single" bond connecting these central atoms to the benzene ring, the unusually small distance of **1 40** A is obtained These results are discussed in relation to acetylene and other compounds

As in the stilbene structure, there are four molecules in the unit cell, and two of these make independent contributions to the asymmetric unit of the crystal The orientations of these two molecules are deduced, and the co-ordinates of the atoms are given The minimum intermolecular approach distances are between 3 5 and 3 6 A for the carbon atoms

REFERENCES

Badger and Bauer 1937 *J Chem Phys* **5**, 599
Ginsburg and Barker 1935 *J Chem Phys* **3**, 668
Herzberg, Patat and Spinks 1934 *Z Phys* **92**, 87
Herzberg, Patat and Verleger 1937 *J Phys Chem* **41**, 123
Mecke 1930 *Z Elektrochem* **36**, 803
Penney and Kynch 1938 *Proc Roy Soc* A, **164**, 409
Robertson 1935*a* *Proc Roy Soc* A, **150**, 348
— 1935*b* *Proc Roy Soc* A, **150**, 106
Robertson and Woodward 1937 *Proc Roy Soc* A, **162**, 568

Relaxation methods applied to engineering problems

II Basic theory, with applications to surveying and to electrical networks, and an extension to gyrostatic systems

By A N Black, B A and R V Southwell, F R S *

(*Received* 22 *June* 1937—*Revised* 24 *November* 1937)

Introduction and summary

1 The "Method of Systematic Relaxation of Constraints" was devised for the determination of stresses in frameworks,—that is in elastic structures having the characteristic that a strained configuration can be specified by attaching values to a *finite* number of co-ordinates Recently it has been extended to continuous systems (e g beams) on the understanding that a finite number of co-ordinates will define a configuration for practical purposes, though not from a mathematical standpoint So far the power of the method has been exhibited only in relation to elastic problems in these its results appear to converge rapidly, judged by a few examples of which the exact solutions were known

The aim of the present paper is fourfold, viz

(1) to prove that relaxation methods, applied systematically to problems of elastic equilibrium, give solutions which converge steadily towards exact results, and hence, by analogy,

(2) to show that they are applicable to any "minimal" problem, e g the adjustment of errors according to the method of least squares,

(3) to notice, as particular examples, the adjustment of errors in level or triangulation surveys, and the partition of electric current in non-inductive networks of conductors, and

(4) to discuss the more difficult problem of an inductive network carrying alternating current This serves as a simple illustration of systems which are governed by equations containing "gyrostatic" or "non-energetic" terms, and which for that reason do not present minimal problems of the usual kind

In regard to (1) it should be mentioned that a proof of convergence was attempted in the original relaxation papers (Southwell 1935 *a*, *b*), but that the arguments there given are open to objection The simple proof now given (§§ 3–4) appears to be satisfactory, and it permits a generalization (2) which greatly widens the scope of the relaxation method The examples mentioned under (3) point to applications in other important fields of engineering, while the problem treated under heading (4) seems to be of interest not only in itself, but especially as indicating potentialities of the relaxation method in relation to problems such as the question of hydrodynamic stability, where similar difficulties are presented by the occurrence of "non-energetic" terms in the governing equations (Chitty and Southwell 1930, III, §§ 12–14)

---

* The idea of applying relaxation methods to the adjustment of errors, together with major credit for the argument in §§ 3–4, should be attributed to Mr Black A joint paper seemed desirable, in view of the close analogy between his problems and those relating to electrical networks which I had studied (R V S)

I On the convergence of the relaxation process

2 The basis of the relaxation method as originally applied to frameworks is the idea that in any specified problem, although it may be difficult to obtain the exact solution, it is easy to calculate the forces which must be applied in order to maintain specified displacements Accordingly we imagine that rigid constraints hold all the joints fixed initially, before the specified forces are applied subsequently, one constraint is "relaxed" in a direction such that the whole or part of its external force is transferred to the framework and through it to adjacent constraints, and the procedure is repeated for other constraints A table of "standard operations", constructed as a first step, enables us to keep account of the displacements and forces entailed by each relaxation Attention is focused on the "residual" forces (i e forces still acting on the constraints), the aim throughout being to reduce these everywhere to zero

Problems of self-straining (Southwell 1935*a*, § 2) can be brought within the scope of the method For when the members of a framework are not so proportioned that they will all fit together without straining, we can calculate forces which will give them such strains that they will fit, and we can start from a "datum configuration" in which these forces are operative and entail (for equilibrium) equal and opposite forces on the constraints Such parts of the latter as do not cancel in the assembled framework may now be regarded as external forces and liquidated by relaxation as before *

3 To show that the relaxation procedure, pursued systematically, gives results converging steadily towards the required solution, it is only necessary to observe that at each step, *if positive work is done on the relaxed constraint*, the total energy of the system (i e elastic strain-energy stored in the frame-

---

* Cf § 6

work *plus* potential energy of the external forces) must be reduced therefore the system must tend continuously towards the required configuration of equilibrium, in which (by a known theorem in Mechanics) the total energy has its minimum value * We shall not arrive at this configuration in any finite number of steps, but we can approach it as closely as we may wish The approach can be accelerated by an introduction of "block relaxations", in which any number of points move together like a rigid body, or of "group relaxations", in which any number move together in some arbitrarily chosen way (Richards 1937 foreword, p 201)

4 The simplest example is provided by a plane *pin-jointed* framework Here, in the notation of previous papers (Southwell 1933, §7) the total strain-energy is given by

$$U = \tfrac{1}{2}\Sigma_m[\Omega_{AB}\{(x_B-x_A)(u_B-u_A)+(y_B-y_A)(v_B-v_A)\}^2], \qquad (1)$$

Σ_m denoting a summation which extends to every member Ω_{AB} is a constant defining the elastic properties of the member which connects A, B, and u_A, v_A are the (small) displacements of the joint A from its position in the unstrained configuration

The total potential energy of the (specified) external forces is given by

$$V = \text{const} - \Sigma_j(\bar{X}_A\ u_A + \bar{Y}_A\ v_A), \qquad (1)$$

Σ_j denoting a summation which extends to every joint of the framework, and $\bar{X}_A$, $\bar{Y}_A$ the components in the x and y directions of the force which is applied at A

Let $\bar{\bar{X}}_A$ denote the "residual force" (§2) which acts on the constraint at A Then initially (when all joints are held fixed) $\bar{\bar{X}}_A = \bar{X}_A$, the specified external force at A, but subsequently (after one or more constraints have been relaxed) a force will also be exerted by the framework This second force (X_A, say) will be equal and opposite to the force exerted by the constraint upon the framework, i e (by Castigliano's first theorem) we have

$$\left.\begin{aligned} X_A &= -\frac{\partial U}{\partial u_A}\\ \text{Therefore, since}\qquad \bar{X}_A &= -\frac{\partial V}{\partial u_A}\quad \text{according to (1),}\\ \text{we may write}\quad \bar{\bar{X}}_A = X_A + \bar{X}_A &= -\frac{\partial}{\partial u_A}(U+V), \end{aligned}\right\} \qquad (2)$$

---

* Kirchhoff's theorem of uniqueness of solution (Southwell 1936*b*, § 14) shows that there is only one configuration of equilibrium, therefore only one absolute minimum of the quantity which is varied in the relaxation process

and the equation $$\frac{\partial}{\partial u_A}(U+V)=0, \tag{ii}$$

which is one condition for a minimum value of $(U+V)$, is then seen to be a condition of equilibrium at A, since it requires that no force shall be left acting on the corresponding constraint

Since $\bar{X}_A$ is independent of the displacements, we have from (2)

$$\left.\begin{aligned} \frac{\partial \bar{\bar{X}}_A}{\partial u_A} &= \frac{\partial X_A}{\partial u_A} = -\frac{\partial^2 U}{\partial u_A^2}, \\ \frac{\partial \bar{\bar{X}}_A}{\partial v_A} &= \frac{\partial X_A}{\partial v_A} = -\frac{\partial^2 U}{\partial v_A\,\partial u_A}, \\ \frac{\partial \bar{\bar{X}}_A}{\partial u_B} &= \frac{\partial X_A}{\partial u_B} = -\frac{\partial^2 U}{\partial u_B\,\partial u_A}, \\ &\text{etc ,} \end{aligned}\right\} \tag{3}$$

and since U is a quadratic function of the displacements, all of the quantities (3) will have constant values independent of the configuration Thus when U is given by (1) we have

$$\left.\begin{aligned} &\frac{\partial^2 U}{\partial u_A^2} = \Sigma_A[\Omega_{AB}(x_B-x_A)^2] = \Sigma_A[(x,x)_{AB}], \\ &\frac{\partial^2 U}{\partial v_A\,\partial u_A} = \Sigma_A[\Omega_{AB}(x_B-x_A)(y_B-y_A)] = \Sigma_A[(x,y)_{AB}], \\ &\frac{\partial^2 U}{\partial u_B\,\partial u_A} = -\Omega_{AB}(x_B-x_A)^2 = -(x,x)_{AB}, \\ &\qquad \text{etc} \end{aligned}\right\} \tag{4}$$

in the notation of "influence coefficients" (Southwell 1935*a*, §12),* Σ_A denoting a summation which extends to every member connected with A Accordingly $\bar{\bar{X}}_A$ is a linear function of the displacements, reducing to $\bar{X}_A$ when all of the displacements are zero, and hence, using (3), we may write

$$\bar{\bar{X}}_A = \bar{X}_A - u_A\frac{\partial^2 U}{\partial u_A^2} - v_A\frac{\partial^2 U}{\partial v_A\,\partial u_A} - u_B\frac{\partial^2 U}{\partial u_B\,\partial u_A} - \quad \text{etc} \tag{5}$$

with corresponding expressions for the other "residual forces"

These expressions hold in respect of any elastic system, and according to

---

* The reciprocal relations $(x,y)=(y,x)$, etc (Southwell 1935*a*, §13) are now seen to be consequences of the commutative property $\dfrac{\partial^2}{\partial u\,\partial v} \equiv \dfrac{\partial^2}{\partial v\,\partial u}$

(5) a residual force $\overline{X}_A$ can be brought to zero by imposing a displacement Δu_A, where

$$\Delta u_A = \overline{X}_A \Big/ \frac{\partial^2 U}{\partial u_A^2} \tag{6}$$

Now because strain-energy is a necessarily positive quantity, the coefficients of u_A^2, u_B^2, etc in U will always be positive, and therefore such differential coefficients as $\partial^2 U/\partial u_A^2$ (For the framework this assertion is exemplified by the first of (4), since the Ω's of necessity are positive) Therefore according to (6) any residual force can be brought to zero by imposing a displacement having the same direction and sense This is the basis of the "relaxation" procedure

Moreover, provided that the displacement is not of such magnitude that the sign of the residual force is changed, it will necessarily involve a *decrease* in the value of $(U+V)$, since in that event

$$\delta(U+V) = -\int \overline{X}_A \, du_A \quad \text{according to (2)}$$

will be negative * The greatest possible decrease for a displacement of given type (u_A) will be obtained by bringing the corresponding force ($\overline{X}_A$) to zero Thus the argument of § 3 is confirmed, together with the theorem there quoted that in the equilibrium configuration $(U+V)$ has a *minimum* value, —not merely a stationary value, which is all that is stated by equations of type (ii) For we have shown that *the value of $(U+V)$ can be reduced so long as any residual force remains on a constraint*, and by a finite amount so long as any residual force is finite hence it follows that the relaxation process can always be continued until $(U+V)$ has been brought so near to its absolute minimum that all residual forces are negligible

II On the wider application of the relaxation method

5 Looking at the method in this way, we see that it is applicable to other problems in engineering and in physics For a framework we have to determine that distribution of joint displacements which (subject to the overriding condition of continuity) entails a minimum value of the total energy, this energy being (by Hooke's law) a quadratic function of the displacements by analogy, whenever we want the conditions for a minimum value

---

* Generalizing the meaning of "force" and "displacement" (Southwell 1936 *b*, §§ 28–30), we can arrive at a similar conclusion in regard to "block relaxations"

of some quadratic function Q of parameters $u_A, u_B, \quad v_A, v_B, \quad$,[*] we can treat those parameters as "displacements" and proceed exactly as before "Joint relaxations" will now be changes made, one at a time, in the values of the parameters, and the partial derivatives of Q with respect to the parameters will measure "forces" which the relaxations impose on the constraints Thus every minimal theorem in mathematical physics provides a fresh application of the relaxation method

U and V now denoting the quadratic and linear parts of Q, we may write in conformity with (2) of § 4

$$\bar{\bar{X}}_A = X_A + \bar{X}_A = -\frac{\partial}{\partial u_A}(U+V), \quad \text{etc},$$

where $\bar{X}_A, \quad$ etc have values independent of $u_A, v_A, \quad$ etc $\bar{\bar{X}}_A$ will again be given by the expression (5), and we have

$$\frac{\partial \bar{\bar{X}}_A}{\partial u_A} = -\frac{\partial^2 U}{\partial u_A^2} = -\frac{\partial^2 Q}{\partial u_A^2}, \quad \text{since } V \text{ is linear,}$$

$$< 0, \text{ by the minimal property of } Q$$

So the same argument as before shows that relaxation processes, applied systematically, will give results which converge steadily towards an exact solution[†]

An exact elastic analogue is obtained if we visualize a framework problem (§ 4) in which $u_A, \quad$ etc are joint displacements, $\bar{X}_A, \quad$ etc are externally applied forces, and $\bar{\bar{X}}_A$ as given by (5) is the total force which comes upon the constraint at A This problem we can treat in the usual way, for equations of type (4) can be used to calculate the effects of joint displacements, and devices such as "block" and "group displacements" (§ 3) can be employed as may be found convenient

Example 1 *Adjustment of errors in a level survey*

6 A simple example is afforded by the problem of adjusting errors in a survey of levels Let $z_A, z_B, \quad$ stand for the heights above datum level of points $A, B, \quad$ for which observations have been made, and let Δ_{AB} stand for the observed difference of level between A and B,—i e more precisely,

---

[*] To preserve the analogy we here divide the parameters into classes, but such grouping is not essential, nor (if it is adopted) is the number of classes in any way restricted

[†] I e to the conditions for an absolute minimum of Q It can usually be proved that Q has only one minimum (cf footnote, p 449)

the observed rise in passing from A to B. On account of errors of observation, the Δ's will not all be consistent with any set of values for z_A, z_B, etc. That is to say, if e_{AB} is the "error" in Δ_{AB}, defined by

$$e_{AB} = \Delta_{AB} - (z_B - z_A), \tag{7}$$

some at least of the e's will be finite whatever values we attach to z_A, z_B, etc. According to the theory of errors, our problem is to determine values such that the sum of the squares of the errors, weighted by reference to considerations which do not concern us here, shall have its minimum value. Initially the z's are unknown, the Δ's are specified.

In the elastic analogue (§ 5), the z's may be regarded as displacements imposed by specified external forces upon a number of yokes or collars (fig. 1), which can slide without friction on a straight bar having the direction Oz, and which are connected by springs having elastic properties proportionate to the "weight factors": our problem is to determine the configuration of equilibrium. Alternatively we may suppose that initially the springs have such lengths that they cannot all be fitted into place without straining, and that e_{AB} is the amount by which the unstrained length (Δ_{AB}) of the spring connecting A and B exceeds the distance ($z_B - z_A$) between those points: then the required configuration is one of self-straining without external forces, which (§ 2) can be brought within the scope of the relaxation method.*

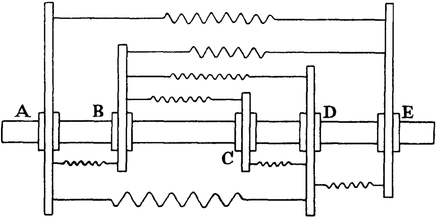


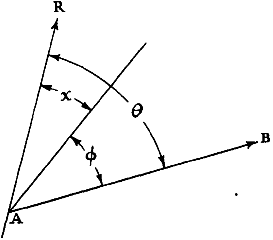
FIG. 1

---

* A numerical example, involving six points of observation, was easily solved by a relaxation process involving twenty six operations.

Example 2 *Adjustment of errors in a triangulation survey*

7 A problem of greater complexity (and as such providing an even more useful application of the relaxation method) is provided by a "triangulation survey" Here the observations made at every point of a network are the angular bearings of other points in relation to the "zero" of the horizontal circle of the theodolite, which zero is given a purely arbitrary orientation Two co-ordinates determine the position of any point, and these can be altered independently in the process of adjustment On account of observational errors no one set of values can be found for the co-ordinates which is entirely consistent with the observations our problem is to assign values to the co-ordinates such that the sum of the squares of the consequent "errors" (weighted by reference to considerations which do not concern us here) shall have its minimum value



Let x_A, y_A and x_B, y_B be values assigned to the co-ordinates of two points A and B of the network, and let the consequent bearing of B at A be θ_{AB}, measured from a fixed reference direction OR As stated above, the bearing of B in relation to A was measured from a reference direction which is unknown If we assign a value χ_A (fig 2) to the unknown angle between this reference direction and the fixed direction OR, the "error" in the observed bearing of B (ϕ_{AB}, say) will be given by

$$e_{AB} = \phi_{AB} - (\theta_{AB} - \chi_A) \tag{8}$$

By assigning a suitable value to χ_A we could make e_{AB} zero, *but the same value must be taken in the similar expressions for* e_{AC}, *etc* , and accordingly χ_A, as well as x_A, y_A, x_B, y_B, is a variable parameter in the minimal problem The ϕ's are specified, and the θ's can be expressed in terms of the x's and y's

8 The "framework analogue" of this problem does not present itself in practice But we can visualize an assemblage of rigid bars, slotted to receive pins which form the joints (fig 3), and sliding smoothly on those pins, and we can suppose that each pin carries a rigid member (A, fig 3) connected by spring constraints to every bar which slides upon that pin We may postulate that each spring constraint (one is indicated diagrammatically in fig 3) tends to preserve a definite angle ϕ between the directions of A and of the bar to which it is connected the angle may in fact have a different value, but if so strain-energy will be stored in the spring Since A is connected with a number of bars all sliding on its pin, although it is free to rotate it cannot (in general) assume an orientation such that none of the springs is strained By Castigliano's second theorem it will assume an orientation χ_A, relative to some fixed direction Ox, such that the total strain-energy has its minimum value and if θ is the orientation, relative to Ox, of any member connected with it, the strain-energy stored in the corresponding spring will be proportional to $\{\phi - (\theta - \chi_A)\}^2$, the constant of proportionality depending on the weight factor, as in § 6 Again the elastic analogue of our problem is a self-strained system

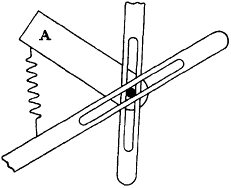


Fig 3

Example 3 *Partition of electric current in a non-inductive network of conductors*

9 Suppose that two nodal points A and B of the network are connected by a conductor of resistance R_{AB}, and let v_A, v_B be the potentials of A and

B By Ohm's law, a current of magnitude $(v_A - v_B)/R_{AB}$ will flow from A to B, so that if I_A, I_B denote the currents flowing towards A and B respectively, then

$$-I_A = I_B = K_{AB}(v_A - v_B),$$

where $K_{AB} = 1/R_{AB}$ Summing for all conductors which are connected with A, we may write

$$\Sigma_A\,[K_{AB}(v_B - v_A)] + \bar{I}_A = 0, \tag{9}$$

where $\bar{I}_A$ stands for the current supplied to A from outside

Now the heat generated in AB will be measured by $K_{AB}(v_A - v_B)^2$, and hence the total generation of heat is given by

$$2U = \Sigma_m[K_{AB}(v_A - v_B)^2], \tag{i}$$

Σ_m denoting a summation extending to every conductor in the network Also, if the current $\bar{I}_A$ is supplied to A from an outside source at datum potential v_0, the energy thereby expended will be measured by $\bar{I}_A(v_A - v_0)$, and hence the total energy expended will be measured by

$$\Sigma_j[\bar{I}_A(v_A - v_0)] = -V \quad \text{(say)} \tag{ii}$$

We see that (9) is a condition for a minimum value of

$$\begin{aligned} Q &= U + V \\ &= \tfrac{1}{2}\,\Sigma_m[K_{AB}(v_A - v_B)^2] + \Sigma_j[\bar{I}_A(v_0 - v_A)], \end{aligned} \tag{10}$$

since it is equivalent to

$$-\frac{\partial Q}{\partial v_A} = -\frac{\partial}{\partial v_A}(U + V) = 0$$

So U is the total strain-energy, V the total potential energy of the external forces, in the elastic analogue which we obtain by interpreting the v's as displacements, the $\bar{I}$'s as external forces, and the K's as "spring constants" or "influence coefficients" Corresponding with the theorem of minimum total energy in the elastic analogue, we can enunciate a minimal theorem relating to the generation of heat by electric currents (the "Joule effect" Jeans 1923, § 355) as follows *In a network of conductors to which specified currents are supplied at two or more nodal points,* the actual distribution of steady currents is such that the total generation of heat*, less *twice the energy expended in supplying the specified currents from a source at datum potential, has its minimum value consistent with the satisfaction of Kirchhoff's second law*

---

* The algebraic sum of the currents must be zero, in order that Kirchhoff's laws may be satisfied

10 The foregoing is to be distinguished from a more familiar theorem relating to the Joule effect (Jeans 1923, § 358) *When a system of steady currents flows through a network of conductors containing batteries of which the electromotive forces are specified, the currents are so distributed that the total generation of heat, less twice the output of energy from the batteries, has its minimum value consistent with the satisfaction of Kirchhoff's first law*

This, in the elastic analogue of § 9, is Castigliano's second theorem For if as before the electric potentials are interpreted as displacements and the K's as "spring constants", and if now the electromotive force E_{AB} of the battery in any link AB of the network is interpreted as the amount $(-\lambda)$ by which the length of the corresponding spring AB exceeds the distance between the points in the datum configuration which it has to connect (cf § 6), then for consistency the current I_{AB} which flows from A to B, and which is measured by $K_{AB}(V_A - V_B + E_{AB})$ must be interpreted as the thrust $(-P)$ to which the spring AB is subjected in a strained configuration, and if $2Q$ is the quantity (total heat generation − twice the energy output from the batteries), i e if

$$2Q = \Sigma_m[I^2/K - 2EI],$$

then in the elastic analogue Q will stand for the quantity

$$\tfrac{1}{2}\Sigma_m\left[\frac{P^2}{K} - 2\lambda P\right]$$
$$= (\text{total strain-energy stored in the elastic system}) - \Sigma_m\,[P\lambda],$$

which is required in Castigliano's second theorem to have a minimum value (Southwell 1923, § 7)

Just as, in that theorem, the variations of Q are restricted by conditions of equilibrium (the analogue of Kirchhoff's first law) and compatibility of displacements is not realized in any but the correct configuration (Southwell 1936*a*, § 2), so in the electrical theorem only the correct distribution of current satisfies Kirchhoff's second law, i e gives a single value for the potential of every nodal point (Jeans 1923, § 346) The theorem of § 9 lends itself more readily to calculation by relaxation methods, and a network of which some or all links contain batteries can be readily treated Starting with the assumption that the whole e m f of each battery is utilized in passing current through the resistance of its own associated "link", we can obtain a datum configuration in which calculable currents enter or leave the network at its nodal points Then we have merely to calculate and superpose the effects of neutralizing currents supplied at the nodal points

III EXTENSION OF THE RELAXATION METHOD TO "GYROSTATIC" SYSTEMS

11 We now consider the more difficult problem in which alternating current flows through a network characterized by self-induction and/or capacity This as relating to sinusoidal current of constant frequency closely resembles the problem of stress-distribution in a plane pin-jointed framework, but the differences are such as to necessitate, in particular instances, a complete revision of the relaxation technique *

Let AB be any link in the network having a vector "admittance" $Y_{AB} = g_{AB} + jb_{AB}$ (the reciprocal of a vector "impedance" $Z = r - jx$), and let

$$V_A = u_A + jv_A, \quad V_B = u_B + jv_B$$

be the vector potentials of the nodal points A, B † Then the vector current flowing from A to B through the link AB is

$$I_{AB} = (g_{AB} + jb_{AB})(V_A - V_B),$$

and hence, if

$$I_A = X_A + jY_A = -\Sigma_A[I_{AB}]$$

is the total vector current flowing *into* A from all links which connect it with other nodal points of the network, we have

$$\left.\begin{aligned} -X_A &= \Sigma_A[g_{AB}(u_A - u_B) - b_{AB}(v_A - v_B)], \\ -Y_A &= \Sigma_A[b_{AB}(u_A - u_B) + g_{AB}(v_A - v_B)], \end{aligned}\right\} \quad (11)$$

Σ_A having the same significance as in preceding sections Moreover, if

$$\bar{I}_A = \bar{X}_A + j\bar{Y}_A$$

stands for the vector current supplied to A from outside, then

$$\overline{\overline{X}}_A = \bar{X}_A + X_A = 0, \quad \overline{\overline{Y}}_A = \bar{Y}_A + Y_A = 0, \quad \text{etc} \quad (12)$$

since there can be no accumulation of current at a nodal point

Equations analogous with (11) and (12) control the joint-displacements of a plane pin-jointed framework Thus we have from (1), (2) and (4) of § 4,

$$\left.\begin{aligned} -X_A &= \frac{\partial U}{\partial u_A} = \Sigma_A[(x,x)_{AB}(u_A - u_B) + (x,y)_{AB}(v_A - v_B)], \\ -Y_A &= \frac{\partial U}{\partial v_A} = \Sigma_A[(y,x)_{AB}(u_A - u_B) + (y,y)_{AB}(v_A - v_B)], \end{aligned}\right\} \quad (13)$$

---

* We are indebted to Professor L N G Filon for calling our attention to this fact

† The notation of this section is based on that of Steinmetz (1900, chap VII), but some new symbols have been introduced with the object of emphasizing the elastic analogue j stands for $\sqrt{(-1)}$

and the conditions of equilibrium can be written in the form of (12) But whereas in the elastic problem Maxwell's reciprocal theorem requires that

$$(x, y) = (y, x), \tag{14}$$

so that $(v_A - v_B)$ in the first of (13) has the same coefficient as $(u_A - u_B)$ in the second, in equations (11) the coefficients, though equal in magnitude, *have opposite signs*, and in consequence, whilst in the elastic problem the equations (12) of equilibrium are conditions for a stationary value of the total potential energy $(U + V)$, which has its minimum value in the required configuration, equations (11) and (12) are conditions for a stationary value of the function

$$\begin{aligned} Q = \tfrac{1}{2}\Sigma_m[g_{AB}(u_B - u_A)^2 - 2b_{AB}(u_B - u_A)(v_B - v_A) - g_{AB}(v_B - v_A)^2] \\ - \Sigma_j[\overline{X}_A \; u_A - \overline{Y}_A \; v_A], \end{aligned} \tag{15}$$

which, in the required configuration, is a minimum as regards variations in the u's, *but a maximum as regards variations in the v's*

Example 4 *Partition of alternating current in an inductive network*

12 No feature of the relaxation process will be altered by reason of the circumstance that Maxwell's relations (14) are not satisfied, and in general, suitably applied,* it will lead quickly to the required solution Occasionally, however, it will fail by giving divergent results, and then a different procedure will be necessary As an example which yields to the standard treatment we shall calculate the amplitude and phase of the current which will pass through the network shown in fig 4 when an alternating potential difference of 1 volt is applied to the nodal points A, E (The analogous problem in relation to a plane framework would be to determine the effects of imposing a relative displacement on two nodal points)

The vector admittances of the different links, calculated for the particular frequency of the applied potential difference, will be taken to have values as under

$$\left.\begin{array}{llll} AB, & 101\ 50 + j\,79\ 12 & BD, & 1\ 78 + j\,6\ 88 \\ AC, & 7\ 58 - j\ \ 8\ 99 & BE, & 2\ 36 + j\,3\ 72 \\ AD, & 12\ 18 - j\ \ 2\ 03 & CD, & 2\ 55 + j\,0\ 55 \\ BC, & 8\ 73 + j\ \ 6\ 72 & CE, & 1\ 78 + j\,3\ 06 \\ & & DE, & 7\ 26 + j\,0\ 73 \end{array}\right\} \tag{16}$$

---

* At every stage the largest of the "residual forces" should be liquidated, and by that operation which affects it more than any other

13 Adopting the procedure suggested at the end of § 10, we first suppose that the vector potential of A is 1 and that B, C, D are at zero potential Then the currents passing from A along the links AB, AC, AD will be given respectively by

$$I_{AB} = (101\ 5+j\,79\ 12), \quad I_{AC} = (7\ 58-j\,8\ 99), \quad I_{AD} = (12\ 18-j\,2\ 03), \quad (1)$$

and no current will flow in any other link of the circuit By addition we deduce that (to make the assumed potentials correct) a current of amount $(121\ 26+j\,68\ 10)$ will have to be supplied to A from outside currents I_{AB}, I_{AC}, I_{AD} will then leave the network at B, C and D respectively

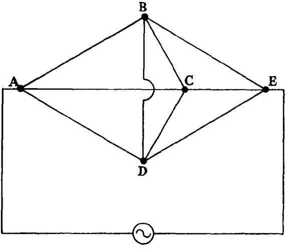


FIG 4

Actually no current passes to or from the network at B, C or D Therefore we must superpose on the assumed potentials those which would result if currents $\bar I_B$, $\bar I_C$, $\bar I_D$, equal respectively to I_{AB}, I_{AC}, I_{AD} as given in (1), were supplied at B, C, D and allowed to leave the network at A and E, the latter points being held at zero potential Writing $\bar I_B = \bar X_B + j\bar Y_B$, etc we thus have initially

$$\left.\begin{aligned} &\bar{\bar{X}}_A = \bar X_A = -121\ 26, \quad \bar{\bar{X}}_B = \bar X_B = 101\ 50, \\ &\qquad\qquad\qquad \bar{\bar{X}}_C = \bar X_C = 7\ 58, \quad \bar{\bar{X}}_D = \bar X_D = 12\ 18, \\ &\bar{\bar{Y}}_A = \bar Y_A = -68\ 10, \quad \bar{\bar{Y}}_B = \bar Y_B = 79\ 12, \\ &\qquad\qquad\qquad \bar{\bar{Y}}_C = \bar Y_C = -8\ 99, \quad \bar{\bar{Y}}_D = \bar Y_D = -2\ 03, \end{aligned}\right\}$$

with

$$\bar{\bar{X}}_E = \bar{\bar{Y}}_E = 0 \quad (17)$$

In the framework analogue these are initial values of the "forces on the constraints", and they have now to be "liquidated" by the imposition of suitable displacements (actually, vector potentials) at B, C and D, *but not at A and E*

14 Next we derive a table of standard operations by calculating the changes in the "residual forces" (typified by $\overline{X}_A$, $\overline{Y}_A$) which result from changes made in any one, *singly*, of the "displacements" (typified by u_A, v_A) As in a normal elastic problem (§ 4), since the "external forces" (typified by X_A, Y_A) are specified and therefore invariant, the partial differentials $\partial\overline{X}_B/\partial u_A$, $\partial\overline{Y}_B/\partial u_A$, are identical with $\partial X_B/\partial u_A$, $\partial Y_B/\partial u_A$, etc According to (11) they will be zero unless B is directly linked with A in the electrical network, and when B is so linked they will have constant values as under

$$\left.\begin{aligned} &\frac{\partial\overline{X}_B}{\partial u_A} = g_{AB} = \frac{\partial\overline{Y}_B}{\partial v_A}, \quad -\frac{\partial\overline{X}_B}{\partial v_A} = b_{AB} = \frac{\partial\overline{Y}_B}{\partial u_A} \\ \text{Similarly} \quad &\frac{\partial\overline{X}_A}{\partial u_A} = \frac{\partial\overline{Y}_A}{\partial v_A} = -\Sigma_A(g_{AB}) = g_{AA}\,(\text{say}), \\ -&\frac{\partial\overline{X}_A}{\partial v_A} = \frac{\partial\overline{Y}_A}{\partial u_A} = -\Sigma_A(b_{AB}) = b_{AA}\,(\text{say}) \end{aligned}\right\} \qquad (18)$$

Here Σ_A stands (as before) for a summation extending to every nodal point which is directly linked with A

As in § 4, each of these differentials gives the effect on some residual force of some unit displacement, and a displacement of different magnitude will have a proportionate effect Table I records the effects of unit operations as deduced from (18) for the example specified in § 12 The operations numbered 7 and 8 are simple examples of "group relaxations"

15 Starting a relaxation table by recording in the first line initial values of $\overline{X}_A$, $\overline{Y}_A$, etc as given in (17), we can now liquidate these "residual" forces in the ordinary way The actual table is not reproduced, but it is summarized in Table II * The progress of the approximation was estimated by the extent to which $\overline{X}_A$, $-\overline{X}_E$ and $\overline{Y}_A$, $-\overline{Y}_E$ had approached equality by thirteen operations the values of $(\overline{X}_A + \overline{X}_E)/(\overline{X}_A - \overline{X}_E)$ and of

$$(\overline{Y}_A + \overline{Y}_E)/(\overline{Y}_A - \overline{Y}_E)$$

were both brought below 1%

---

* The relaxation was performed for us by Mr R J Atkinson

TABLE I

Column number		1	2	3			4	5	6	
Number and nature of operation	$\bar{X}_A$	$\bar{X}_B$	$\bar{X}_C$	$\bar{X}_D$	$\bar{X}_E$	$\bar{Y}_A$	$\bar{Y}_B$	$\bar{Y}_C$	$\bar{Y}_D$	$\bar{Y}_E$
(1) $u_B = 1$	101 50	**−114 37**	8 73	1 78	2 36	79 12	−96 44	6 72	6 88	3 72
(2) $u_C = 1$	7 58	8 73	**−20 64**	2 55	1 78	−8 99	6 72	−1 34	0 55	3 06
(3) $u_D = 1$	12 18	1 78	2 55	**−23 77**	7 26	−2 03	6 88	0 55	−6 13	0 73
(4) $v_B = 1$	−79 12	96 44	−6 72	−6 88	−3 72	101 50	**−114 37**	8 73	1 78	2 36
(5) $v_C = 1$	8 99	−6 72	1 34	−0 55	−3 06	7 58	8 73	**−20 64**	2 55	1 78
(6) $v_D = 1$	2 03	−6 88	−0 55	6 13	−0 73	12 18	1 78	2 55	**−23 77**	7 26
(7) $u_B = 1,\ v_B = 1$	22 38	−17 93	2 01	−5 10	−1 36	180 62	**−210 81**	15 45	8 66	6 08
(8) $u_B = 1,\ v_B = -1$	180 62	**−210 81**	15 45	8 66	6 08	−22 38	17 93	−2 01	5 10	1 36

TABLE II

Column number		1	2	3			4	5	6	
Operation and multiplier	$\bar{X}_A$	$\bar{X}_B$	$\bar{X}_C$	$\bar{X}_D$	$\bar{X}_E$	$\bar{Y}_A$	$\bar{Y}_B$	$\bar{Y}_C$	$\bar{Y}_D$	$\bar{Y}_E$
Initial values	−121 26	101 50	7 58	12 18	0	−68 10	79 12	−8 99	−2 03	0
(1) × 0 9482	96 24	−108 45	8 28	1 69	2 24	75 02	−91 44	6 37	6 52	3 53
(2) × 0 8503	6 45	−2 79	−17 55	2 17	1 51	−7 64	5 71	−1 14	0 47	2 60
(3) × 0 6904	8 41	1 23	1 76	−16 41	5 01	−1 40	4 75	0 38	−4 23	0 50
(4) × (−0 0289)	2 29	−2 79	0 19	0 20	0 11	−2 93	3 31	−0 25	−0 05	−0 07
(5) × (−0 1744)	−1 57	1 17	−0 23	0 10	0 53	−1 32	−1 52	3 60	−0 44	−0 31
(6) × 0 0109	0 02	−0 07	−0 01	0 07	−0 01	0 13	0-02	0 03	−0 26	0-08
Final values	−9 42	0 01	0 02	0	9 39	−6 24	−0 05	0	−0 02	6 33

From the operations and multipliers given in the first column we can deduce the vector voltages Rejecting the fourth decimal place as unreliable, we have

$$\left.\begin{aligned} v_B &= 0\ 948 - j\,0\ 029, \\ v_C &= 0\ 850 - j\,0\ 174, \\ v_D &= 0\ 690 + j\,0\ 011 \end{aligned}\right\} \tag{19}$$

From the last line of Table II we see that the total current flowing through the network from A to E is

$$9\ 4 + j\,6\ 3 \tag{20}$$

A MODIFICATION OF THE STANDARD METHOD

16 When (as in this example) the residual forces decrease with satisfactory rapidity, we may regard the standard procedure as justified by its results, for they vanish in the exact solution, which is unique * But examples can be constructed in which, as relaxation proceeds, the residual forces are found to *increase*, and for use in such cases we now describe a modification of the standard procedure which, like that procedure as applied to normal elastic problems, can be shown to lead always to convergent results

17 The required configuration is one in which, for each of a number of specified nodal points,

$$\overline{\overline{X}} = \overline{\overline{Y}} = 0, \tag{i}$$

$\overline{\overline{X}}, \overline{\overline{Y}}$ being defined by (11) and (12) for a typical point A If now $\Sigma_{(j)}$ stands for a summation extending to those nodal points, *but to those alone*, for which (i) is to be satisfied, and if

$$2F = \Sigma_{(j)}[\overline{\overline{X}}^2 + \overline{\overline{Y}}^2], \tag{21}$$

then F will vanish and so (being necessarily positive) will attain its minimum value in the required configuration Like the total potential energy $(U + V)$ in a normal elastic problem, *it can always be reduced so long as any "residual force" ($\overline{\overline{X}}$ or $\overline{\overline{Y}}$) remains "unliquidated"*

We can thus treat F exactly as in § 5 we treated the function denoted by Q Corresponding with the residual forces of that section we have "quasi-forces" $\boldsymbol{X}_A$, $\boldsymbol{Y}_A$, etc given in terms of F by expressions of the types

$$\boldsymbol{X}_A = -\frac{\partial F}{\partial u_A}, \quad \boldsymbol{Y}_A = -\frac{\partial F}{\partial v_A}, \tag{22}$$

---

* We have, in effect, to solve a system of simultaneous linear equations, equal in number to the number of the unknowns

which are analogous with (2), and corresponding with the influence coefficients defined in equations (3) and (4) we have "quasi-influence coefficients" given by such quantities as

$$\frac{\partial^2 F}{\partial u_A^2} = -\frac{\partial X_A}{\partial u_A}, \quad \frac{\partial^2 F}{\partial u_A \partial v_A} = -\frac{\partial X_A}{\partial v_A} = -\frac{\partial Y_A}{\partial v_A}, \quad \frac{\partial^2 F}{\partial v_A^2} = -\frac{\partial Y_A}{\partial v_A}, \tag{23}$$

which are analogous with (3) The latter we can insert in a table of standard operations, analogous with Table I, and we can start a relaxation table, analogous with Table II, by inserting in the first line initial values of the X's and Y's as given by (22) when all the u's and v's are zero The relaxation procedure will be the same as before, and its convergence may be established by an argument exactly similar to that of §§ 3 and 4, since by definition (23) the "quasi-influence coefficients" satisfy reciprocal relations of the type of (14)

18 According to (21)

$$\frac{\partial F}{\partial u_A} = \Sigma_{(j)}\left[\bar{X}\frac{\partial \bar{X}}{\partial u_A} + \bar{Y}\frac{\partial \bar{Y}}{\partial u_A}\right], \quad \frac{\partial F}{\partial v_A} = \Sigma_{(j)}\left[\bar{X}\frac{\partial \bar{X}}{\partial v_A} + \bar{Y}\frac{\partial \bar{Y}}{\partial v_A}\right], \tag{24}$$

and the differentials in these expressions have constant values typified by (18) of § 14 Hence

$$\left.\begin{aligned} \frac{\partial^2 F}{\partial u_A^2} = \Sigma_{(j)}\left[\left(\frac{\partial \bar{X}}{\partial u_A}\right)^2 + \left(\frac{\partial \bar{Y}}{\partial u_A}\right)^2\right], \quad \frac{\partial^2 F}{\partial v_A^2} = \Sigma_{(j)}\left[\left(\frac{\partial \bar{X}}{\partial v_A}\right)^2 + \left(\frac{\partial \bar{Y}}{\partial v_A}\right)^2\right], \\ \frac{\partial^2 F}{\partial u_A \partial v_B} = \Sigma_{(j)}\left[\frac{\partial \bar{X}}{\partial u_A}\frac{\partial \bar{X}}{\partial v_B} + \frac{\partial \bar{Y}}{\partial u_A}\frac{\partial \bar{Y}}{\partial v_B}\right], \\ \text{etc} \end{aligned}\right\} \tag{25}$$

Combined with (22) and (23), these expressions show that the "quasi forces" are linear in the u's and v's and that the "quasi-influence coefficients have constant values Thus they justify the assumptions which were made (tacitly) in § 17

19 Using the formulae (25) we can deduce the new table of standard operations (Table III) from an operations table of the normal kind (Table I) In the latter (q v) columns and lines are so arranged that the same number relates to a "corresponding" force and displacement (Southwell 1936*b*, § 28) thus line 2 relates to a displacement u_C, and column 2 to the residual force $\bar{X}_C$ Let the new table be arranged in conformity, so that (for example) line 2 relates to u_C and column 2 to X_C, and suppose that we want to calculate $\partial X_C/\partial u_B$, the entry appropriate to line 1, column 2 (or, since reciprocal

relations are satisfied, to line 2, column 1) Observing that in Table I the entries in lines 1 and 2 of the column appropriate to any force $\overline{X}$ are the values of $\partial\overline{X}/\partial u_B$ and $\partial\overline{X}/\partial u_C$ respectively, we deduce from (25) that for this purpose we must sum the products of such entries for all columns of Table I which relate to forces included within the summation $\Sigma_{(j)}$ Such columns are distinguished by heavy vertical rulings in Table I in Table III nothing is gained by retaining other columns, which accordingly have been left blank

Proceeding as described, we have as the appropriate entry in line 1, column 2 and in line 2, column 1 of Table III

$$\begin{aligned}\frac{\partial X_C}{\partial u_B} = \frac{\partial X_B}{\partial u_C} &= +114\ 37 \times 8\ 73 + 8\ 73 \times 20\ 64 - 1\ 78 \times 2\ 55 \\ &\qquad + 96\ 44 \times 6\ 72 + 6\ 72 \times 1\ 34 - 6\ 88 \times 0\ 55, \\ &= +1827\ 40, \end{aligned} \tag{1}$$

and the other entries are calculated similarly In our example the work is shortened by the circumstance that not only relations of type (14) are satisfied, but also relations typified by

$$\frac{\partial X_B}{\partial u_C} = \frac{\partial Y_B}{\partial v_C}, \quad \frac{\partial X_B}{\partial v_C} = -\frac{\partial Y_B}{\partial u_C}, \quad \frac{\partial X_C}{\partial v_C} = \frac{\partial Y_C}{\partial u_C} = 0 \tag{26}$$

TABLE III

Column number		1	2	3			4	5	6	
Number and nature of operation		X_B	X_C	X_D			Y_B	Y_C	Y_D	
(1) $u_B = 1$	—	−22553 04	1827 40	925 60	—	—	0	183 79	−474 91	—
(2) $u_C = 1$	—	1827 40	−555 98	55 58	—	—	−183 79	0	37 61	—
(3) $u_D = 1$	—	925 60	55 58	−659 90	—	—	474 91	−37 61	0	—
(4) $v_B = 1$	—	0	−183 79	474 91	—	—	−22553 04	1827 40	925 60	—
(5) $v_C = 1$	—	183 79	0	−37 61	—	—	1827 40	−555 98	55 58	-
(6) $v_D = 1$	—	−474 91	37 61	0	—	—	925 60	55 58	−659 90	—

20 To start the new relaxation table we must insert in its first line the initial values of the X's and Y's We have from (22) and (24)

$$X_B\,(\text{initial}) = -\Sigma_{(j)}\left[\overline{X}\frac{\partial\overline{X}}{\partial u_B} + \overline{Y}\frac{\partial\overline{Y}}{\partial u_B}\right], \tag{27}$$

with corresponding expressions for the other "quasi-forces" so because the column numberings of Tables I and II conform, if (for example) we want the entry appropriate to column 2 of the new relaxation table, then we are concerned with the first line (initial "forces") of Table II and with line 2 of Table I Multiplying together the two numbers which appear in any particular column, we have to sum the resulting products Thus in our example the appropriate entry is given by

$$\begin{aligned} X_C\,(\text{initial}) &= -8\ 73\times 101\ 50+20\ 64\times 7\ 58-2\ 55\times 12\ 18 \\ &\qquad -6\ 72\times 79\ 12-1\ 34\times 8\ 99+0\ 55\times 2\ 03, \\ &= -1303\ 32 \end{aligned} \tag{ii}$$

The other entries are found similarly

The new relaxation table has not been reproduced, since the procedure required to liquidate these initial X's and Y's is in every way normal Applied to the particular example of § 12, fourteen operations sufficed to give the same solution as was found by an unmodified procedure in §§ 14–15 *

CONCLUSION

21 The relaxation procedure is in effect a method for obtaining approximate solutions of systems of linear simultaneous equations (The number of independent equations must be equal to the number of the unknowns, in order that the exact solution may be unique) In previous applications (i e to problems of elastic equilibrium) the equations have been conditions for an absolute minimum of a certain quadratic function of the variables, necessarily positive here, in the last example treated, they are conditions for a stationary value of a certain quadratic function, but that value is not an absolute minimum since the function can take either sign By a modification of the standard procedure we have made it applicable to problems of this less restricted class, and in so doing we have *not* made use of the fact that in the particular example treated (which concerns a "gyrostatic" system) certain skew-symmetrical relations govern the coefficients in the several equations Accordingly the procedure as modified (§§ 16–20) will apply without restriction to any system of linear simultaneous equations

REFERENCES

Chitty, L and Southwell, R V 1930 *Philos Trans* A, **229**, 205–53, and *Aero Res Ctee R and M* 1200

Jeans, J H 1923 "The Mathematical Theory of Electricity and Magnetism" Cambridge University Press, 4th ed

---

* The relaxation was performed for us by Mr D G Christopherson.

Richards, J C 1937 *J Inst C E* **5**, 196–216
Southwell, R V 1923 *Phil Mag* **45**, 193–212, and *Aero Res Ctee R and M* 821
— 1933 *Proc Roy Soc* A, **139**, 475–507
— 1935*a* *Proc Roy Soc* A, **151**, 56–95
— 1935*b* *Proc Roy Soc* A, **153**, 41–76
— 1936*a* *Proc Roy Soc* A, **154**, 4–21
— 1936*b* "Introduction to the Theory of Elasticity" Oxford University Press
Steinmetz, C P 1900 "Alternating Current Phenomena" McGraw Publishing Co

---

The reflexion coefficients of ionospheric regions

By E V Appleton, F R S and J H Piddington,* M Sc, B E
Cavendish Laboratory, Cambridge

(*Received* 17 *November* 1937)

In a recent communication† to the Society, Watson Watt, Wilkins and Bowen (1937) have described the results of an ionospheric investigation conducted by means of radio wave reflexion which they have interpreted as demonstrating the existence of four or more highly reflecting strata in the lower atmosphere, about 10 km above ground level These results and their interpretation are naturally of great interest to those whose field of investigation has been the ionospheric regions at greater heights, for, as the authors of (I) state, their own findings "must modify the detailed discussion of ionospheric soundings effected with sounding waves which have had to traverse these highly reflecting strata" Moreover, if the high reflexion coefficients estimated for these low-lying strata are correct, even as regards order of magnitude only, it is immediately obvious that our ideas concerning the travel of wireless waves to great distances require radical revision, for, since 1926, it has been customary to regard such communication as generally being effected either by way of Region E in the case of long waves, or by way of Region F in the case of the shorter waves It can, further, be asserted that current explanations of such phenomena are in such close agreement with the results of what is now known of the structure of the ionosphere at levels of 90 km and above,

* Walter and Eliza Hall Research Fellow

† We shall, for brevity, refer to this paper as (I) throughout

that any evidence not in harmony with these accepted relations must be accorded the closest scrutiny

The experimental observations of Colwell and Friend (1936) and of the authors of (I) (1937) concerning the return of radio waves from low levels in the atmosphere are of the greatest scientific importance and, immediately they were announced, investigations were begun here in Cambridge to see if we could confirm them It was desired especially to ascertain in this connexion how far the existence of such lowlying reflecting strata influenced the interpretation of the vast amount of observational work already carried out in various parts of the world on radio reflexion at higher levels

The results of our own experiments, and the interpretation concerning the ionospheric importance of low-level reflexions which we have felt obliged to give them, were different from those of the above-mentioned authors so that, on the publication of (I), we have examined afresh the arguments adduced in it which have led its authors to certain of their conclusions The account of our own work is therefore prefaced by a brief discussion of previous work

THE EVIDENCE FOR THE EXISTENCE OF HIGHLY REFLECTING STRATA IN THE LOWER ATMOSPHERE

The most complete account of the return of radio echoes from the lower atmosphere is given in (I) and the discussion in this section is therefore devoted to it The authors of (I)

(*a*) have confirmed the occurrence of echoes returned from Region D These were noted by Appleton (1927) as occurring occasionally in 1927,

(*b*) have found there is another reflecting region at 45–50 km with a reflexion coefficient* of 0 3 for 6 Mc /sec waves (We may refer to this as Region C), and

(*c*) have found that there is a series of discrete reflecting strata, persisting without substantial change of level for several days, at such heights as 8 5, 9 3, 10 3, 10 75 and 13 5 km , each with a reflexion coefficient of 0 7 for 6–12 Mc /sec waves (We may refer to this as Region B)

Since, from our own observations, we are unable to confirm the magnitude of the reflexion coefficients stated in (*c*) above we must examine afresh the arguments presented by the authors of (I) in this connexion. We need only take the case of Region B (i.e. the Region Z of (I)).

---

* The reflexion coefficient here refers to the ratio of the reflected and incident electric vectors

The argument which leads to the evaluation of the reflexion coefficients of the low-lying strata proceeds in this way. From a snap-picture of a series of echoes, four or more series of multiple reflexions are identified and, from the amplitude decay among the members of a selected series, the reflexion coefficient of the stratum responsible for the series is calculated. For example, since the fifth echo of a series is, in amplitude, 0.2 of the first, the effective reflexion coefficient is taken as $(0.2)^{\frac{1}{4}}$* (i.e. 0.67). The effective reflexion coefficient corresponding to each series is evaluated independently as if reflexion by, or transmission through, any strata below had no effect. This argument raises the following points.

(*a*) Since we must allow for the spacial attenuation of the waves (a doubly reflected wave has, for example, travelled twice the distance of a singly reflected one) the reflexion coefficient in a case where the fifth order echo is 0.2 of the first order is given by

$$\frac{B_5}{B_1} = \left(\frac{\rho}{2}\right)^4 = 0.2, \tag{1}$$

i.e.

$$\rho = 1.37 \tag{2}$$

Thus the argument leads to the conclusion that the reflexion coefficient of a single stratum is greater than unity. If further allowance were made for the imperfect reflexion of 6 Mc./sec. waves at the ground (for which the reflexion coefficient ρ' is of the order of 0.9) it will be seen that the Region B reflexion coefficient would come out higher still.

(*b*) Even if we accept the conclusion of the authors of (I) that the lowest reflecting stratum in Region B has a reflexion coefficient of 0.7 it is easy to see that, for the effective reflexion coefficient (measured as the ratio of the incident and reflected wave vectors at a point below the lowest reflecting stratum) of the remaining strata to be of the same value, it is necessary for the true reflexion coefficient of the second stratum to be unity. This would mean that substantially no radio energy reaches Regions E and F, so that they would play no really significant part in the propagation of radio waves.

There are other arguments of a more general character which are difficult to reconcile with the permanent existence of highly reflecting strata at fixed distances from the ground. It is to be expected, for example, that such strata, if they exist, would be fixed relatively to the levels of fixed density, pressure or temperature, and yet it is known that such levels vary from day to day Yet the conclusion drawn from the experimental evidence in (I) is that the four strata retain the same heights for several days independently of the physical characteristics of the atmosphere where they are formed Again, it is easy to show that a stratum which would reflect radio waves copiously at levels of 10 km must have an electrical conductivity of the order of, at least, 10^7 e s u Yet the general electrical conductivity at these levels, as measured in balloon flights, is of the order of 10^{-4} e s u It does, however, appear possible, as pointed out by Appleton (1937), to reconcile the important observations described in (I) with previous work, and also with the results of the experiments described below if we assume that the echoes returned from low levels are not reflected from layers at all, but are signals of very low intensity indeed, scattered from sporadic low-lying patches of ionization * If such an assumption is entertained we must abandon the detailed analysis of echoes into correlated groups and regard the gradual decay of amplitude peaks with increase of echo-delay as indicating the successive reception of echoes from scattering centres at different distances The echoes of greater delay, for example, might be caused by scattering centres not immediately overhead

---

* There appears to be a misprint in (I) (p. 186), where $(0.2)^{\frac{1}{5}}$ is written and $(0.2)^{\frac{1}{4}}$ is intended.

EXPERIMENTAL OBSERVATIONS

In order to look for confirmation of the existence of highly reflecting strata in the lower atmosphere a pulse sender and receiver of the required characteristics were built and erected, the sender in a hut on the Solar Physics Observatory site, Cambridge,† and the receiver in a private house 7 km away

(a) *Pulse sender*

This was capable of giving 2 to 3 kW output power on a fixed frequency of 8 8 Mc /sec The pulse duration was variable but could be made as low as 30 μsec The aerial system employed was a half-wave dipole at optimum height for vertical radiation, and was fed from the sender by a matched feeder-line Radiation was therefore strongest in an upward vertical direction

---

* The agency of ionization is here invoked following the general argument of (I) But if our conclusions concerning the order of magnitude of the reflexion coefficients of these scattering patches is correct, the possibility of reflexion from other types of atmospheric discontinuities must not be excluded

† We are greatly indebted to Professor F J M Stratton for hospitality in this connexion

(b) *Pulse receiver*

This was of superheterodyne type with wide frequency band-pass characteristics The output dropped to half value, as tested by a signal generator, when the set was detuned by 33 kc /sec Transmission of the pulses through the receiver lengthened them to 50 μsec , which means that a pulse deflected the base line of the oscillograph over a part of the scale corresponding to the delay of an echo reflected from an equivalent height of 7 km Thus echoes of this magnitude of semi-path and of appreciable intensity would be expected to show as clear from the ground signal

Using a signal generator the pulse to noise ratio was found to be 5 to 1 for a pulse input of 10 μV Due to the good receiving site this noise was mainly of internal (set) and not external origin Assuming an aerial equivalent height of 4 m ,* we may thus say that pulses of minimum field strength of 2 or 3 μV/m would be detectable As at the sending station, a half-wave dipole situated about a quarter of a wave-length above ground was used at the receiving station Due to the low horizontal transmission and reception the receiver ground ray was small In case the reflexions for which we were searching had a particular polarization the receiving aerial was placed first at right angles and then parallel to the sending aerial

(c) *Accessory equipment*

The usual cathode-ray tube oscillograph equipment was used and calls for no special comment except to remark that the time-base scale was very open, echoes of equivalent height differences of 16 km being separated by 1 cm on the screen

(d) *Limits of sensitivity expressed in terms of equivalent reflexion coefficient*

Since the relatively high frequency used penetrated normal Region E at all times, estimates of the relation between reflexion coefficient and echo amplitude were made using Region F reflexions Using these it is possible to express the intensity of a particular echo in terms of an equivalent reflexion coefficient, such coefficient being unity in the case for an ideal perfectly reflecting stratum situated at a height corresponding to the delay time of the echo in question

---

* In this assumption a liberal allowance has been made for losses in the transmission from aerial to receiver

This calibration was effected as follows The receiver gain control was calibrated with a signal generator The amplitudes of the ground ray (G), first Region F reflexion (F') and second Region F reflexion (F'') were measured We then have $G/F' = n_1$ and $F'/F'' = n_2$ Then $\rho = 2/n_2$ and $F' = G/n_1$ If however ρ were raised to unity, F' would increase in the ratio $1/\rho$ (i e $n_2/2$) so that we should have $F' = n_2 G/2n_1$ This quantity was found, and from its magnitude, allowing for spacial attenuation, it was calculated that a layer of unity equivalent reflexion coefficient at 100 km would give a first order echo of $3G$, that a similar layer at 10 km would give a first order echo of $30G$, and so on Using the calibrated volume control it was possible to estimate the magnitude of the echo (expressed as a fraction of the ground wave at maximum workable gain) which was just detectable This was found to be $\frac{1}{320}$, so that it should have been possible to detect a first-order echo from a layer at 10 km when the equivalent reflexion coefficient was of the order of 0 0001 *

It should be noted that the above definition of equivalent reflexion coefficient may be used to denote the intensity of an echo returned from a small scattering centre and not from a layer at all The coefficient in this case gives us the ratio of the echo actually obtained to that which would be obtained if there were a perfectly reflecting stratum at the same distance as the scattering centre

EXPERIMENTAL RESULTS

It may be stated at once that no stratum at 10 km height with a reflexion coefficient of the order of 0 7 has been detected by us during three months of observation We must thus conclude, in accordance with our interpretation of the results of (I) as due to atmospheric scattering patches, that the effective reflexion coefficient of such patches must be less than 0 0001 They would therefore play no substantial part in ionospheric alternation or refraction, so that our ideas concerning the normal processes of wireless transmission are left undisturbed The detection of scattered echoes from lower intensity than the limit stated above obviously requires a higher-powered sender, since the receiver equipment is already operating at full sensitivity Such a sender is now being built

---

* This order of magnitude for the upper limit of ρ may be checked in another way Using formula (3), p 475, and allowing for the effect of ground reflexion at both sender and receiver, it may be shown that a 2 kW sender would produce echoes of the estimated minimum detectable field strength (i e 2 μV/m) if reflected from a centre, the effective reflexion coefficient of which is 0 0001

Through the courtesy of Mr Watson Watt we have been able to test our own receiver in conjunction with the higher-powered sender used by the authors of (I) In this way we have found that, using the method of calibration outlined above, B Region echoes were detected, with 6 Mc /sec waves, of amplitude corresponding to a reflexion coefficient of 0 00007

In the course of our work at Cambridge it has been found that, after the normally recognized regions of the ionosphere (Regions E and F) the reflecting centres next in importance as regards magnitude of echo production are transient in character Such echoes were noted by Appleton, Naismith and Ingram (1936) as coming from above the 100 km level, as lasting for a few seconds and as occurring by night and day * Using the same apparatus described above we have made a study of these transient echoes, paying particular attention to measurements of effective reflexion coefficient and of equivalent height of reflexion

In fig 1 is shown a plot of the relative numbers of these transient echoes as a function of equivalent height, such numbers representing the number of echoes obtained from successive equal ranges of height in a series of observations made during the day It will be seen that the echoes come most frequently from a height of about 110 km Other characteristics of these echoes may be summarized briefly as follows

(*a*) The echoes can be roughly grouped into two series, one group consisting of members lasting a fraction of a second up to one or two seconds, and the other of members lasting several seconds up to some minutes

(*b*) The equivalent reflexion coefficient may be as high as 0 05, though that of those most frequently observed lies between 0 0005 and 0 005

(*c*) They are observed both by day and by night, and a nocturnal graph corresponding to the day-time curve in fig 1 is very similar in character

There can be little doubt that these observations indicate the temporary existence of marked scattering centres in the ionosphere which form and disappear throughout the whole of the twenty-four hours There is a marked resemblance between fig 1 and the corresponding curve for the frequency of occurrence of auroral heights, but detailed consideration of this and other matters is being deferred for a later communication The results however indicate, without further evidence, the entry into the atmosphere, by day and night, of an agency which produces bursts of ionization of considerable intensity

---

* These characteristics have also recently been noted by Mr T L Eckersley, *Nature*, 13 November 1937, who has also succeeded in getting photographic registration of the received echoes (See also fig 10 of (I))

It is possible to estimate the minimum number of electrons required in a scattering cloud to give transient echoes of the type we have described. A cloud of 10^{16} electrons, for example, concentrated within a region of linear dimensions small compared with a wave-length, would give an effective reflexion coefficient of the order of 0·001 as observed. If the cloud were of greater dimensions a greater total number of electrons would be required to give the same effective reflexion coefficient.

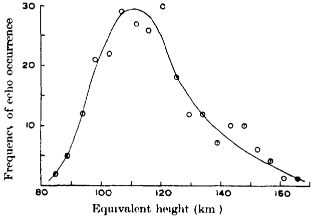


FIG. 1

A corresponding word may be added concerning the number of electrons in a small cloud of given effective reflexion coefficient ρ situated at the 10 km. level. Our experiments show that ρ is less than 0·0001. Let us suppose, then, that it is 0·00002. The number of electrons necessary to give this value of ρ would be $3 \cdot 6 \times 10^{13}$. The corresponding number of ions of molecular mass required would be 2×10^{18}.

The pulse observations described above have been supplemented by observations and enquiries concerning the reflexion of ultra-short waves such as are used in the B.B.C. Television Service since the authors of (I) point out that they have been able to detect Region B reflexions, very frequent and of disturbing amplitude, using frequencies as high as 40 Mc./sec. Our enquiries and our own observations show that no general and sustained reception of "ghost" images, such as would be caused by strong echoes from low-lying strata, has been observed in the area within 90 km. of the sending station. A careful watch was maintained throughout the summer of 1937 on a television receiver situated in St John's College and very occasionally ghost images were seen. These corresponded to various delay times, a typical example being 8 μsec. corresponding to a shift of 1 in on a 10-in picture Since the distance between Alexandra Palace and Cambridge is 69 km this would correspond to an equivalent height of about 10 km for the scattering centre if situated midway and in the plane of propagation This is of the same order of magnitude of Region B levels quoted in (I) A number of the ghost images noted were, however, accompanied by a fluttering of the intensity of the whole television picture, indicating rapid motion of the scattering centre In this way a substantial number of the ghost images noted were found to be due to aeroplanes, and we are not at all sure that any of the cases noted were due to reflexion from atmospheric scattering centres *

In any case it is easy to show, as follows, that the influence of a naturally-occurring scattering centre would be expected to be small in television reception at Cambridge The field strength E of an echo reflected at a low height from a station of power P, at distance r, is given by the formula

$$E \text{ (volts per metre)} = \frac{0\ 3\rho\sqrt{P(\text{kW})}}{r(\text{km})}, \tag{3}$$

where ρ is the effective reflexion coefficient Now the value of P is 3 kW, so that if we take r as 69 km and ρ as 0 00007 the value of E is found to be 0 5 μV per metre The influence of additive interference by way of ground reflexion may, on occasion, give signals of double this value, but even then the field strength would be much smaller than that due to the direct transmission Since, however, our estimate of ρ is probably much too large for the high frequencies we are considering it is very unlikely that B-Region echoes will give rise to sustained interference at any reception point

ACKNOWLEDGEMENT

This work has been carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research

SUMMARY

Experiments on the radio sounding of the ionosphere are described which give no support to the supposed existence of permanent highly reflecting regions at heights of 10 km It is considered possible that the

---

* We are much indebted to Mr Edwards of Pye Radio Ltd for collaboration in these observations

echoes of short delay noted by previous workers are due to atmospheric scattering patches of very low effective reflexion coefficient It is found that, after the normally recognized regions of the ionosphere, the radio reflecting centres in the atmosphere next in order of importance as regards wireless echo production are transitory bursts of ionization which occur at equivalent heights of from 80 to 160 km Such bursts of ionization are due to some cosmic agency which is usually effective throughout the whole of the day and night

REFERENCES

Appleton 1927 *Union Int Radioteleg Sci Pap Gen Assembly Washington*, October, **1**, Pt I

— 1937 *Electrician*, 29 January, p 143

Appleton, Naismith and Ingram 1936 *Phil Trans Roy Soc* A, **236**, 254

Colwell and Friend 1936 *Nature, Lond*, **137**, 782

Watt Wilkins and Bowen 1937 *Proc Roy Soc* A, **161**, 181

The spectrum of turbulence

By G I Taylor, F R S

(*Received* 1 *December* 1937)

When a prism is set up in the path of a beam of white light it analyses the time variation of electric intensity at a point into its harmonic components and separates them into a spectrum Since the velocity of light for all wave-lengths is the same, the time variation analysis is exactly equivalent to a harmonic analysis of the space variation of electric intensity along the beam In a recent paper Mr Simmons (Simmons and Salter 1938) has shown how the time variation in velocity at a field point in a turbulent air stream can be analysed into a spectrum In the present paper it is proposed to discuss the connexion between the spectrum of turbulence, measured at a fixed point, and the correlation between *simultaneous* values of velocity measured at two points

If u, the component at a fixed point of turbulent motion in the direction of the main stream in a wind tunnel, is resolved into harmonic components the mean value of u^2 may be regarded as being the sum of contributions from all frequencies If $\overline{u^2}\,F(n)\,dn$ is the contribution from frequencies between n and $n+dn$, then

$$\int_0^\infty F(n)\,dn = 1 \tag{1}$$

If $F(n)$ is plotted against n the diagram so produced is a form of the spectrum curve

The proof that u^2 may be regarded as being the sum of contributions from all the harmonic components has been given by Rayleigh, using a form of Parseval's theorem If $u = \phi(t)$ and

$$\left.\begin{aligned} I_1 &= \frac{1}{\pi}\int_{-\infty}^{+\infty}\phi(t)\cos\kappa t\,dt, \\ I_2 &= \frac{1}{\pi}\int_{-\infty}^{+\infty}\phi(t)\sin\kappa t\,dt \end{aligned}\right\} \tag{2}$$

Rayleigh showed that

$$\int_{-\infty}^{+\infty}[\phi(t)]^2\,dt = \pi\int_0^\infty (I_1^2+I_2^2)\,d\kappa, \tag{3}$$

or if $\kappa = 2\pi n$, so that n represents the number of cycles per second,

$$\int_{-\infty}^{+\infty}[\phi(t)]^2\,dt = 2\pi^2\int_0^\infty (I_1^2+I_2^2)\,dn \tag{4}$$

If the integrals on the right hand side of (2) and the left-hand side of (3) are taken over a long time T instead of infinity the left-hand side of (3) is $T\overline{u^2}$, so that

$$u^2 = 2\pi^2\int_0^\infty \underset{T\to\infty}{\mathrm{Lt}}\left(\frac{I_1^2+I_2^2}{T}\right)dn \tag{5}$$

The quantity $2\pi^2 \underset{T\to\infty}{\mathrm{Lt}}\left(\frac{I_1^2+I_2^2}{T}\right)$ is therefore the contribution to $\overline{u^2}$ which arises from the components of frequency between n and $n+dn$, i e

$$2\pi^2 \underset{T\to\infty}{\mathrm{Lt}}\left(\frac{I_1^2+I_2^2}{T}\right) = F(n) \tag{6}$$

CONNEXION BETWEEN SPECTRUM CURVE AND CORRELATION CURVE

Now consider two cases (*a*) where the variation in u is due to eddies of small extent which are carried by a wind stream of velocity U past the fixed point, (*b*) the variation is due to large eddies carried in the wind stream In case (*a*) the fluctuations at the fixed point will be much more rapid than they are in case (*b*) The spectrum analysis in case (*b*) will therefore show greater values of $F(n)$ for small values of n than in case (*a*)

It is clear when the eddies are large the correlation R_x between simultaneous values of u at distance x apart must fall away with increasing x more slowly than when the eddies are small One may therefore anticipate that when the (R_x, x) curve has a small spread in the x co-ordinate the $F(n)$ curve will extend to large values of n and vice versa

If the velocity of the air stream which carries the eddies is very much greater than the turbulent velocity, one may assume that the sequence of changes in u at the fixed point are simply due to the passage of an unchanging pattern of turbulent motion over the point, i e one may assume that

$$u = \phi(t) = \phi\left(\frac{x}{U}\right), \tag{7}$$

where x is measured upstream at time $t = 0$ from the fixed point where u is measured In the limit when $u/U \to 0$ (7) is certainly true Assuming that (7) is still true when u/U is small but not zero, R_x is defined as

$$R_x = \frac{\overline{\phi(t)\,\phi\left(t+\frac{x}{U}\right)}}{\overline{u^2}} \tag{8}$$

We now introduce another expression analogous to (3) It can be shown that*

$$\int_{-\infty}^{+\infty} \phi(t)\,\phi\left(t+\frac{x}{U}\right) dt = 2\pi^2 \int_0^\infty (I_1^2 + I_2^2) \cos\frac{2\pi n x}{U}\, du, \tag{9}$$

where I_1 and I_2 have the same meaning as in (3)

Substituting for $I_1^2 + I_2^2$ from (6), (9) becomes

$$\frac{\overline{\phi(t)\,\phi\left(t+\frac{x}{U}\right)}}{\overline{u^2}} = \int_0^\infty F(n) \cos\frac{2\pi n x}{U}\, dn, \tag{10}$$

hence from (8)

$$R_x = \int_0^\infty F(n) \cos\frac{2\pi n x}{U}\, dn \tag{11}$$

It will be noticed that the form of (11) is very similar to that of the Fourier

---

[*] This formula can be deduced from the theorem 9 09 given on p 70 of Norbert Wiener's *The Fourier Integral*, Camb Univ Press, 1933

integral The Fourier integral theorem is usually expressed by the pair of formulae

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(\mu)\cos\mu x\, d\mu, \tag{12}$$

$$g(\mu) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f(x)\cos\mu x\, dx \tag{13}$$

When $f(x)$ is symmetrical, so that $f(x) = f(-x)$, (12) and (13) may be written

$$f(x) = \sqrt{\frac{2}{\pi}}\int_{0}^{\infty} g(\mu)\cos\mu x\, d\mu, \tag{14}$$

$$g(\mu) = \sqrt{\frac{2}{\pi}}\int_{0}^{\infty} f(x)\cos\mu x\, dx \tag{15}$$

Comparing (11) and (14) it will be seen that if

$$\mu = \frac{2\pi n}{U}, \quad f(x) = R_x, \quad g(\mu) = \frac{UF(n)}{2\sqrt{2\pi}},$$

then (11) and (14) are identical

Making these substitutions in (15), the following expression is found for $F(n)$

$$F(n) = \frac{4}{U}\int_{0}^{\infty} R_x \cos\frac{2\pi n x}{U}\, dx \tag{16}$$

It seems therefore that R_x and $\dfrac{UF(n)}{2\sqrt{2\pi}}$ are Fourier transforms of one another

If $F(n)$ is observed we can calculate R_x using (11), and if R_x is observed we can calculate the spectrum curve $F(n)$ using (16)

COMPARISON WITH OBSERVATION

Measurements have been made by Mr L F G Simmons of R_x and of $F(n)$ at a point 6 ft 10 in from a turbulence-producing grid with a mesh 3×3 in at wind speeds $U = 15, 20, 25, 30$ and 35 ft /sec It was found that except very close to $x = 0$, R_x is nearly independent of U within that range * When R_x is independent of U it will be seen from (16) that $UF(n)$ must be a function of n/U

Accordingly Mr Simmons' measurements of $F(n)$ for all values of U have

---

* A similar result has been obtained by Dryden, N A C A report 581, 1937

been plotted on the same diagram (figs 1 and 1 *a*), in which the ordinates are $UF(n)$ and the abscissae are n/U The fact that the points fall so closely on one curve is very satisfactory evidence that the measurements of $F(n)$ are accurate

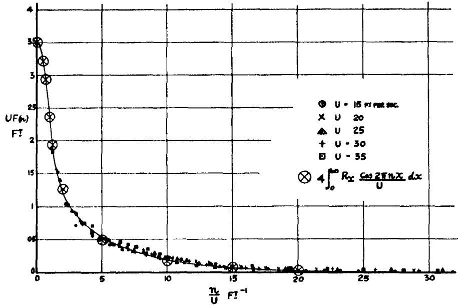


Fig 1

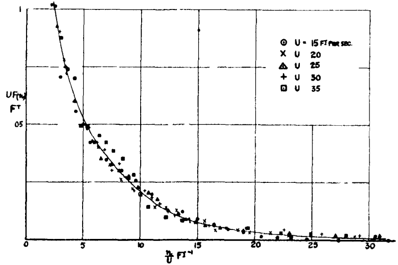


Fig 1 *a*

Mr Simmons' measurements of R_x are shown in fig 2 The values of $F(n)$ calculated using the measured values of R_x in (16) are shown in fig 1, but as the points corresponding with the lower part of fig 1 are rather close together an enlarged version is shown in fig 1*a* The values of R_x calculated using the measured values of $F(n)$ in (11) are shown in fig 2

It will be seen that the agreement in both cases is good

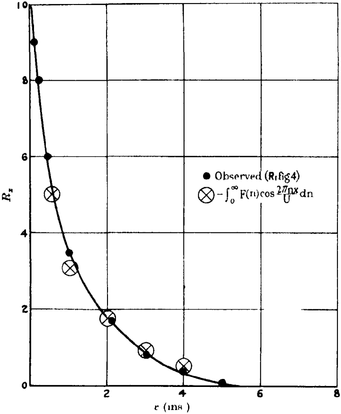


FIG 2

It has been seen that the points in fig 1 seem to fall on one curve, as is predicted by (16), when it is assumed that R_x is independent of U On the other hand it is known that the curvature of the R_x curve at its vertex is not independent of U This curvature is defined (Taylor 1935) by means of a length λ, where

$$\frac{1}{\lambda^2} = 2 \operatorname*{Lt}_{x \to 0} \left(\frac{1 - R_x}{x^2}\right) \tag{17}$$

If the turbulence is isotropic experiments show that λ is proportional to $U^{-\frac{1}{2}}$

To find what effect this variation in λ with U may be expected to produce in the $F(n)$ curve, λ may be expressed in terms of $F(n)$ When n is small $\cos \frac{2\pi n x}{U}$ may be replaced in (11) by $1 - \frac{2\pi^2 x^2 n^2}{U^2}$ Hence

$$\frac{1}{\lambda^2} = \frac{4\pi^2}{U^2}\int_0^\infty n^2 F(n)\, dn \tag{18}$$

Since (18) can be written in the form

$$\frac{1}{\lambda^2} = 4\pi^2 \int_0^\infty \frac{n^2}{U^2} U F(n) \frac{dn}{U}, \tag{19}$$

λ must be independent of U if the $\{UF(n), n/U\}$ curve is independent of U This deduction is inconsistent with the observed fact that λ is proportional to $U^{-\frac{1}{2}}$

The explanation of this apparent discrepancy is that the value of $\int_0^\infty n^2 F(n)\, dn$ depends chiefly on the values of $F(n)$ for large values of n, i e on the parts of figs 1 and 1 *a* where the points are so close to the axis that variations in their height above it are hardly visible In fig 3 the vertical scale of $UF(n)$ has been enlarged very greatly It will be seen that above $n/U = 16$ the $UF(n)$ curves separate, that for $U = 15$ ft /sec falling below those for 20 and 35 ft /sec

CALCULATION OF λ FROM THE SPECTRUM CURVE

To determine λ from the spectrum curve the integral (19) must be evaluated using the values of $UF(n)$ taken from figs 1 and 3 It is instructive to tabulate the contributions to this integral which arise from various ranges of n/U These are set forth in Table I, where they are expressed in ft -sec units It will be seen that when $U = 35$ about half of the integral is due to components for which $n/U > 30$, in spite of the fact that the highest

TABLE I CONTRIBUTIONS TO $\int_0^\infty \frac{n^2}{U^2} F(n)\, dn$ EXPRESSED IN FT -SEC UNITS

n/U	$U = 15$	$U = 20$	$U = 35$
0–16	24 0	24 0	24 0
17–30	7 7	19 0	23 9
31–∞	0	15 5	44 4
Total	31 7	58 5	92 3

value of $F(n)$ in this range is only $\frac{1}{200}$ of its maximum value (namely 0 35 when $n = 0$) The value of λ in feet is found by inserting the numbers given at the foot of Table I for the integral in (19) Thus when $U = 35$ ft /sec $\lambda = (4\pi^2 \times 92\ 3)^{-\frac{1}{2}} = 0\ 00165$ ft $= 0\ 50$ cm The values of λ so calculated are given in column 2, Table II

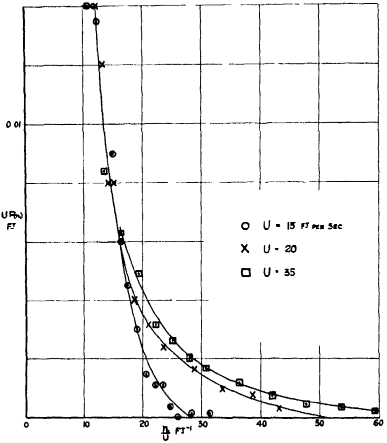


FIG 3

VALUE OF $F(n)$ AT $n = 0$

Putting $x = 0$ in (16)

$$U[F(n)]_{n=0} = 4\int_0^\infty R_x\,dx$$

By numerical integration of the measured R_x curve in fig 2 it is found that

$$\int_0^\infty R_x dx = 1\ 07 \text{ in } = 0\ 089 \text{ ft},$$

so that, when $n = 0$,

$$U[F(n)]_{n=0} = 4 \times 0\ 089 = 0\ 35 \text{ ft}$$

This upper limit is marked in fig 1

PROOF THAT TURBULENCE IS ISOTROPIC

Though the theory and measurements so far discussed do not involve any assumption as to whether or not the turbulence is statistically isotropic, yet since the theory of isotropic turbulence has been discussed so completely it is worth while to describe measurements which prove that the turbulence was in fact isotropic

It has been shown by Karmán (1937) that when turbulence is isotropic there is a definite relationship between the correlation curves R_x and R_y Kármán defines two correlation functions R_1 and R_2 R_1 is the correlation between components of velocity along the line AB, where A and B are the points at which the velocities are measured, R_2 is the correlation between components at right angles to AB In isotropic turbulence R_1 and R_2 are functions of r only where r is the length AB When correlation measurements are made in a wind tunnel by means of a hot wire, only the component parallel to the length of the tunnel produces any appreciable effect on the hot wire If therefore A and B are situated on a line parallel to the mean wind stream the correlation R_x is identical with Karmán's R_1 If A and B are situated on a line perpendicular to the stream the correlation R_y is identical with Kármán's R_2

Kármán's relationship between R_1 and R_2, namely

$$R_2 = R_1 + \tfrac{1}{2}r\frac{dR_1}{dr}, \tag{20}$$

is therefore a relationship between the correlations R_x and R_y which have been measured The measured values of R_x or R_1 are given in fig 2, and they are repeated in fig 4 To this curve the (negative) values of $\frac{1}{2}r(dR_1/dr)$ are added and the calculated values of R_y or R_2 thus obtained are shown in fig 4 The values of R_y measured by Mr Simmons at 6 ft 10 in behind a 3×3 in grid are also shown in fig 4 It will be seen that Kármán's relationship (20) is very well verified, and it may fairly be concluded that the turbulence at 6 ft 10 in behind a 3 × 3 in grid in a wind tunnel is isotropic

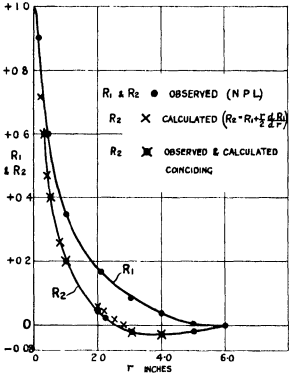


FIG 4

CALCULATION OF λ FROM MEASURED RATE OF DISSIPATION OF ENERGY

In the case of isotropic turbulence λ can be found by measuring the rate of dissipation of energy This can be found by measuring the rate of decay of the mean kinetic energy of turbulent motion

Fig 5 shows the measured values of U/u', $(u' = \sqrt{\overline{u^2}})$, in the air stream in which $F(n)$ and R_x were measured It will be seen that in this case U/u' increases linearly* with x, the distance down stream from the grid I have shown (Taylor 1935) that U/u' increases linearly with x when λ satisfies the following relationship

$$\frac{\lambda}{M} = A\sqrt{\frac{\nu}{u'M}}, \tag{21}$$

---

* This is not a general law Cases where U/u' does not increase linearly have been observed

where M is the mesh size of the grid producing the turbulence Conversely, when U/u' increases linearly with x, λ must be related to u' by the equation (21) By measuring the slope of the line which passes through the observed points in fig 5, I find that A in (21) is 2 12 At the point where the spectrum measurements were made, 6 ft 10 in from the grid, $U/u' = 33\ 5$ Since $M = 3$ in $= 7\ 62$ cm and $\nu = 0\ 148$, (21) becomes

$$\lambda = 2\ 12\sqrt{\frac{33\ 5\times 0\ 148\times 7\ 62}{U}}\ \text{cm}$$

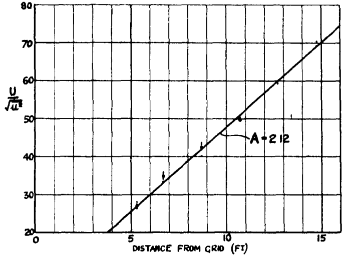


FIG 5

The values given in column 3, Table II, are calculated from this formula Comparing the values of λ calculated from the measured dissipation (column 3) with those calculated from the measured spectrum curves using equation (18) (column 2), it will be seen that the agreement is fairly good

TABLE II VALUES OF λ

U ft /sec	λ calculated from spectrum curves	λ calculated from observed dissipation
15	0 86 cm	0 61 cm
20	0 63	0 53
35	0 50	0 40

CORRELATION MEASURED WITH BAND FILTER CIRCUITS

Recently Dryden (1937) has made measurements of R_x using various band filter circuits in his amplifier The action of the band filter is to cut out all disturbances except those whose frequencies lie between certain limits By supplying truly sinusoidal disturbances of known frequency and amplitude to the filter circuit the characteristic curve showing the response of the circuit to unit input were obtained If $\phi(n)$ is the ratio of $\overline{u^2}$ measured with the filter to $\overline{u^2}$ measured without it, the characteristic curve $\{\phi(n), n\}$ for one of Dryden's circuits which passes frequencies between 250 and 500 cycles is shown in fig 6 The measured values of R_x using this filter circuit and with wind speed $U = 20$ ft /sec are shown in fig 7

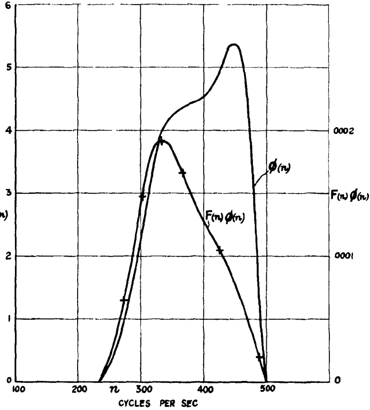


FIG 6

We have already seen how R_x is related to $F(n)$ It is clear that the same relationship will still hold when the filter circuit is inserted, but $F(n)$ must now be replaced by

$$\frac{F(n)\,\phi(n)}{\int_0^\infty F(n)\,\phi(n)\,dn}$$

If $R_{x\phi}$ is the value of R_x measured with the band filter circuit ϕ, the formula analogous to (11) is

$$R_{x\phi} = \frac{\int_0^\infty F(n)\,\phi(n)\cos\frac{2\pi nx}{U}\,dn}{\int_0^\infty F(n)\,\phi(n)\,dn} \tag{22}$$

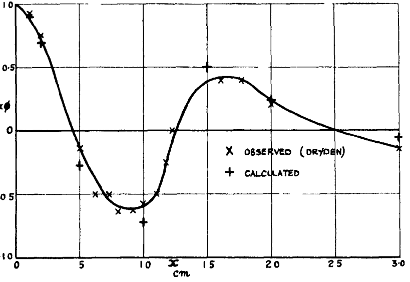


FIG 7

Before this formula can be used in comparison with Dryden's observed values of $R_{x\phi}$ it is necessary to find $F(n)$ Dryden has not measured $F(n)$ except very roughly, but he has measured R_x in the same air stream and without the filter circuit His measurements are shown in fig 8 To calculate $F(n)$ we may use the expression (16) with the observed R_x The values of $\frac{U}{4}F(n)$ so found is also shown in fig 8 Taking the values of $F(n)$ from a smooth curve through the calculated points, the values of $F(n)\,\phi(n)$ together can be found for any given values of U and m by multiplying the ordinates of the $F(n)$ and $\phi(n)$ curves The values of $F(n)\,\phi(n)$ at $U = 20$ ft /sec ($= 610$ cm /sec) found in this way are shown in fig 6

Using a series of values of x, the values of $R_{x\phi}$ have been calculated by numerical integration of (22) The values so calculated are shown in Table III, and are marked in fig 7 It will be seen that the agreement with Dryden's observations is very good This agreement provides additional evidence in favour of the main thesis of this paper that R_x and $\frac{U}{2\sqrt{2\pi}}F(n)$ are Fourier transforms of one another, so that each can be predicted when the other has been measured

TABLE III—CALCULATED VALUES OF $R_{x\phi}$

x (cm)	0.1	0.2	0.5	1.0	1.5	2.0	3.0
$R_{x\phi}$	+0.91	+0.79	−0.27	−0.72	+0.51	+0.23	−0.06

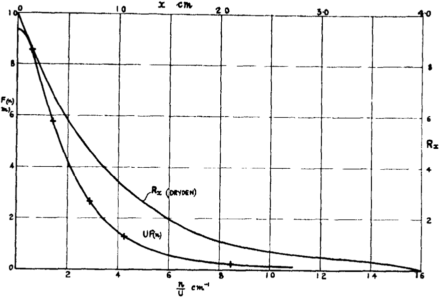


FIG 8

BEARING OF SPECTRUM MEASUREMENTS ON THEORY OF DISSIPATION

The fact that the $\{UF(n), n/U\}$ curve is independent of U over nearly the whole range indicates that the turbulent flow at a fixed point behind a regular grid is similar, so far as the main features of the flow are concerned, at all speeds On the other hand the fact that small quantities of very high frequency disturbances appear, and increase as the speed increases, seems to confirm the view frequently put forward by the author that the dissipation of energy is due chiefly to the formation of very small regions where the vorticity is very high Apart from these very small regions the turbulence behind a grid is similar at all speeds

SUMMARY

It is shown that a definite connexion exists between the spectrum of the time variation in wind at a fixed point in a wind stream and the curve of correlation between the wind variations at two fixed points The spectrum curve and the correlation curve are, in fact, Fourier transforms of one another

As an example of the use of this relationship the spectrum of turbulence in an American wind tunnel was calculated from measurements of correlation by Dryden In some further experiments Dryden modified this spectrum by inserting a filter circuit and then measured the correlation with this filter in circuit The modified spectrum is here calculated from the filter characteristics and the Fourier transform theorem is used to calculate the modified correlation curve The agreement with Dryden's measurements is very good indeed

The paper ends with some remarks on the bearing of the spectrum measurements on the theory of dissipation of energy in turbulent flow

REFERENCES

Dryden, Schubauer Mock and Skramstad 1937 "Measurements of intensity and scale of wind tunnel turbulence and their relation to the critical Reynolds number of spheres" National Adv Comm Aeronautics, No 581 (Referred to as Dryden 1937)

Kármán, T de 1937 *J Aero Sci* **4** 131

Simmons and Salter 1938 *Proc Roy Soc* A (In the press)

Taylor, G I 1935 *Proc Roy Soc* A, **151**, 421

Structure of stretched rubber

By C J B Clews and F Schoszberger
University of Vienna

(*Communicated by Sir William Bragg, O M, P R S — Received* 24 *July* 1937)

[Plate 7]

Investigations of the micellar structure of fibre substances have given rise to two theories The older theory (Meyer and Mark 1930, Mark 1932, Siefriz 1934, Meyer 1930, and Nageli 1928) considers the micelles as separate crystallites, between which lie the intermicellar spaces The micelles consist of "Hauptvalenzketten"* bound together along their length by homeopolar bonds and in the transverse direction by van der Waals' forces, the intermicellar binding being also attributed to van der Waals' forces The original model suggested in work published by K H Meyer (1930), for cellulose, depicts the micelles arranged like bricks in a wall (fig 1), and doubtless this is the simplest explanation of the X-ray results But it is difficult to understand how such an arrangement can give a micellar structure its peculiar mechanical properties, and further how it is possible, when both inter- and intramicellar cohesion are attributed to the same type of force, to cause by swelling experiments an enlargement of the intermicellar spaces, while the "Hauptvalenzketten" remain unaffected

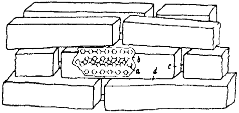


Fig 1 Micellar structure after Meyer *a*, Hauptvalenzketten, *b*, intramicellar regions, *c*, intermicellar holes, *d*, intermicellar long spaces

An alternative theory has been put forward by O Gerngross, K Herrmann and W Abitz (1930), W T Astbury (1933), A Frey-Wyssling

---

* Long chain molecules of high molecular weight

(1936) and E Guth and S Rogowin (1936) These authors suppose that a given "Hauptvalenzkette" is not confined to a single crystalline region but may stretch through more such regions In general, the arrangement of the neighbouring chains will be truly lattice-like, but a chain may lie at too great a distance from its neighbours or not lie exactly parallel to them, so that the structure as a whole will show statistically distributed spaces In fig 2 ordered crystalline regions may be distinguished (drawn in thick line), but their significance is physically different from that of the crystallites of the Meyer model They are not self-contained units, the whole system is linked together due to the "Hauptvalenzketten" extending beyond a single micelle Astbury considers that in a substance of high molecular weight of a type capable of swelling that part which produces the X-ray spectrum is the concentration centre of a complicated network of thread like molecules He draws an analogy between micellar structure and the secondary structure of Zwicky He suggests that it is possible that micellar systems, which are characterized by a mixture of perfection and imperfection, are the counterpart in compounds of high molecular weight of the well known mosaic structure of the more familiar crystals Frey-Wyssling is of the opinion that the micelles, growing together, enclose lens-shaped spaces running parallel to the fibre axis Between these intermicellar spaces are small rod-shaped regions of undistorted lattice, which are the so-called micelles of the earlier work (fig 3) In this figure, which gives a pictorial representation of Frey's theory, the statistically distributed hollow spaces are shown black, some of these are enclosed in undistorted crystalline regions A lamellar structure consisting of superimposed monomolecular layers suggested by O L Sponsler and W H Dore (1930) has been shown to be untenable from the work on double refraction by Baas-Becking and Gallliher (1931)

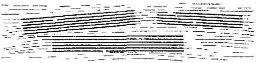


FIG 2 Micellar structure after Guth and Rogowin

In order to investigate further the micellar structure the effect of "higher orientation ', which up till now has been little used in such cases, was studied (Mark and Kratky 1937) By this is meant an ordering of the crystalline regions so that they are parallel not only to the long axis, but also in a second crystallographic direction Fig 4 represents a cross-section through a fibre system, the short lines being sections of the single chains Regions which are exactly ordered others not completely ordered and hollow spaces are shown The *a* and *c* axes lie in quite different directions in different parts of the fibre but the *b* axis (fibre axis) is always in approximately the same direction, i e in our diagram, perpendicular to the plane of the paper With the appearance of "higher orientation", the direction of the "Hauptvalenzketten" remains as before, but a parallel ordering of the *a* and *c* axes occurs (The lines will be parallel to one another) The effect was first observed by Mark and v Susich (1928) for stretched rubber the intensity of the interference pattern alters in a characteristic way

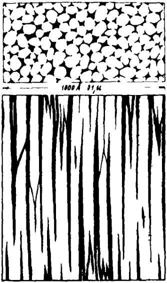


Fig 3 Cross section and longitudinal section through an intermicellar system (A Frey Wyssling)

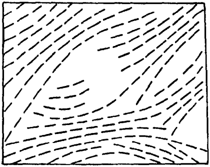


Fig 4 Cross section of a fibre system (Mark and Kratky)

according to whether the radiation is directed parallel or perpendicular to the direction of stretching With the radiation perpendicular to the direction of stretching, the spot (200) is most intense, (210) is weaker, while (002) and (012) are feeble, the order being reversed when the radiation parallel to the direction of stretching, (002) is very strong, (012) less so and (200) and (210) are quite weak From such diagrams Mark and v Susich have recalculated the rhombic elementary cell of rubber to be $a = 12\ 3$, $b = 8\ 1$ (fibre axis) and $c = 8\ 3$ A

A tentative structure with eight isoprene residues in the unit cell and based on a series of isoprene chains has been put forward. From measurements of the breadths of interference patterns Hengstenberg (1928) estimates the dimensions of the rubber micelle to be of the order 600 × 500 × 150 A. Recent work of Sauter (1937) is in substantial agreement with this structure although the dimensions of the unit cell are found to be slightly larger: $a = 12·60 \pm 0·05$ A, $c = 8·91 \pm 0·05$ A, $b = 8·20 \pm 0·05$ A. It has been further suggested by Meyer and Mark (1928) that the elementary cell for rubber must be shorter than the length of a single molecule.

Other examples of "higher orientation" are shown by the cell wall of *Valona ventricosa*, where it appears with growth, and by mechanically worked cellophane. From the results of a number of experiments (Mark and v. Susich 1929, Burgeni and Kratky 1929, Eggert and Luft 1930, Hess and Trogus 1930) it was concluded that the lath-shaped micelles become oriented by rolling so that their longest and second longest edges lie in the plane of rolling. Mark and Kratky (1937) consider that sliding in the lattice or rotation of the single chains cannot give rise to "higher orientation". In order to explain the latter on a basis of interlocking chains and statistically distributed spaces, the micelles must be lath-shaped and must possess a certain individuality. This will occur if the statistically distributed spaces appear more frequently in the direction of two crystallographic axes perpendicular to the fibre axis. Then the small regions of undistorted lattice will automatically become bound together by lattice forces into lath-like forms and will cease to be isolated entities.

It is now interesting to discuss whether the building up of such lath-like aggregates is a characteristic of naturally occurring conditions, or whether the "Hauptvalenzketten" themselves are able to build up such micelles from solution. It has already been observed that hydrated cellulose films obtained from xanthogenate solutions (Burgeni and Kratky 1929) show "higher orientation", but it might be argued that these solutions were extremely concentrated and that the micelles themselves had not been broken up.

We have therefore investigated a series of very dilute solutions of a rubber soluble with difficulty, made by drying the latex (braun-gelb) in benzene, chloroform and carbon tetrachloride, these solvents being chosen with a view to examining a possible influence on "higher orientation". The rubber only partially dissolved after shaking for 1–2 weeks on a machine, so that the objection that the micelles have not been split up can still be made. For comparison, solutions of an easily soluble rubber (first latex crepe) were also made in the same solvents. To make the films actually

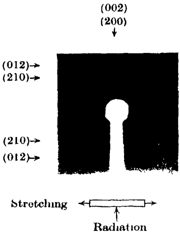


Fig 5 *a*

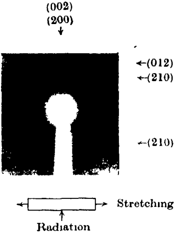


Fig 5 *b*

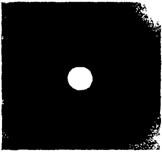


Fig 6 *a*

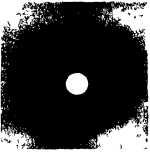


Fig 6 *b*

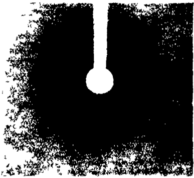


Fig 7 *a*

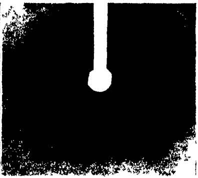


Fig 7 *b*

employed, the dilute solutions were poured on to a glass plate or mercury surface, when the solvent evaporated and left a thin film of rubber, by repeating the process several times the films could be made about 0·1 mm thick They were then dusted with rice powder after which they could be stripped from the underlying surface The X-ray photographs were taken on a fibre camera, the plate-specimen distance being 35 mm and the collimator 0 8 mm in diameter The rubber film under investigation was held in a special extension apparatus

The X-ray tube was a Siemens Strukturrohre with a copper antikathode and Ni filter, running at 45 kV and 15 mA Under these conditions the exposure time was of the order of 6 hr

All the films derived from these solutions show "higher orientation" for X-rays on extreme extension to about 700 % (see figs 5 and 6, plate 7) It is interesting to note that the effect also appears with a film of vulcanized rubber (fig 7, plate 7) The results of this work establish that the capacity for building up lath-shaped micelles in the way described is an intrinsic property of the "Hauptvalenzketten" themselves

The picture of micelle structure here developed is intermediate between the two formerly proposed theories The micelles are of lath-shaped form but probably not closed on all sides, and possess, on account of their shape and of their transverse bindings, a definite individuality, even though the "Hauptvalenzketten" themselves extend beyond the sphere of single micelles

We have pleasure in thanking Professor H Mark and Dr O Kratky for their continued advice and assistance, and their interest in the investigation One of us (C J B C) is indebted to Professor H Mark for extending to him the hospitality of his Institute and to the University of London for a travelling studentship which made the work possible

REFERENCES

Astbury, W T 1933 *Trans Faraday Soc* **29**, 193
Bass Becking and Galliher 1931 *J Phys Chem* **35**, 467
Burgeni, A and Kratky, O 1929 *Z phys Chem* B, **4**, 401
Eggert, J and Luft, F 1930 *Z phys Chem* B, **7**, 468
Frey Wyssling, A 1936 *Protoplasma*, **25**, 262
Gerngross, O, Herrmann, K and Abitz, W 1930 *Biochem Z* **228**, 409
Guth, E and Rogowin, S 1936 *S B Akad Wiss Wien*, **145**, 531
Hengstenberg 1928 *Z Kristallogr* **66**, 637
Hess, K and Trogus, C 1930 *Z phys Chem* B, **9**, 169
Mark, H 1932 "Physik und Chemie der Cellulose" Berlin

Mark, H and Kratky, O 1937 *Z phys Chem* B, **36**, 129
Mark, H and Susich, G v 1928 *Kolloidzschr* **46**, 11
— — 1929 *Z phys Chem* **4**, 431
Meyer, K H 1930 *Kolloidzschr* **53**, 8
Meyer, K H and Mark, H 1928 *Ber dtsch chem Ges* **61**, 1939
— — 1930 ' Aufbau der hochpolymeren organischen Naturstoffe " Leipzig
Nägeli, C 1928 "Die Micellarthoerie " Leipzig
Sauter, E 1937 *Z phys Chem* B, **36**, 405
Siefriz, W 1934 *Protoplasma*, **21**, 129
Sponsler, O L and Dore, W H 1930 *Zellulose Chem* **11**, 186

DESCRIPTION OF PLATE 7

FIG 5*a* and 5*b* X ray photographs of a crepe rubber film obtained from a chloroform solution Higher orientation is clearly shown * 700 % stretching Radiation perpendicular and parallel to plane of film respectively

FIG 6*a* and 6*b* X ray interference diagrams for a film of raw rubber (braun gelb) from a benzene solution 700 % stretching

FIG 7*a* and 7*b* Photographs showing higher orientation of a film of vulcanized rubber 800 % stretching

* Cf figs 1*a* and 1*b* of Mark and v Susich (1928)

Statistical mechanics of the adsorption of gases at solid surfaces

BY F J WILKINS

Research Department, Imperial Chemical Industries, Billingham

(*Communicated by R H Fowler, F R S —Received* 20 *September* 1937)

The object of this paper is to develop a statistical mechanical treatment of adsorption of gases on plane solid surfaces using the method of partition functions Several authors have approached this problem by way of the Boltzmann equation but without success In fact Kruyt and Moddermann (1930), who discuss this work in some detail, conclude that the equation cannot be satisfactorily applied to the problem of gaseous adsorption at a solid surface and suggest that this is because the forces causing adsorption are of a chemical nature

Recently, Fowler (1936*a*) has developed the Langmuir isotherm using essentially Langmuir's assumption of chemisorption on an adsorbing surface similar to a chequer board containing a fixed number of points of adsorption, and further the assumption that if a molecule is adsorbed on one point it does not influence adsorption on contiguous points This further assumption is relaxed in later papers (Fowler 1936*b*, Peierls 1936) The assumption made by Langmuir that adsorption forces are chemical in nature is not always true, as it is likely that the forces operative are in many cases of the van der Waals' type His other assumptions are unnecessarily restrictive in view of the considerable amount of evidence on lateral mobility in the adsorbed layer and also of the certainty that molecules adsorbed contiguously suffer the normal molecular interactions of the van der Waals' type

A rather more general treatment is attempted therefore in this paper

The derivation of an adsorption isotherm from the Boltzmann equation itself is readily achieved, provided a suitable model of the adsorbed phase is chosen In the first place it must be assumed that the potential throughout the adsorption volume is uniform In the second place, the adsorbed layer is assumed to be a fluid phase similar to a compressed perfect gas The adsorption volume can be considered therefore to be a part of the system in which the gas comes under the influence of an external field of force Fowler (1936*a*, p 63) gives the Boltzmann equation as

$$\bar{a}_1/\bar{a}_2 = w_1 e^{-\epsilon_1/kT}/w_2 e^{-\epsilon_2/kT}$$

Applying this to the problem under consideration we write $\bar{a}_1$ and $\bar{a}_2$ as the average number of molecules in equal elements of translational phase space which have the total energies ϵ_1 and ϵ_2 w_1 and w_2 are the corresponding weights for the internal energies of the molecules For equal volumes of physical space accessible to the systems, in which their dynamical state is the same, integrating over all possible momenta, we obtain the result

$$n_1/n_2 = w_1 e^{-W_1/kT}/w_2 e^{-W_2/kT},$$

where W_1 and W_2 are the potential energies of the molecules in the two portions of physical space This reduces to

$$C_a/C_g = w_2/w_1 e^{\phi/RT},$$

where C_a and C_g are the number of g mol of gas in equal volumes of the adsorbed and gaseous phases respectively and ϕ is the adsorption potential Therefore

$$x = p\frac{\alpha d}{RT}\frac{w_2}{w_1}e^{\phi/RT}, \tag{1}$$

where x is the total number of g mol adsorbed per unit area at pressure p and temperature T, α is the fraction of the area available for adsorption and d is the thickness of the adsorbed layer

It is important to note that x is not the number of molecules adsorbed as it is usually determined, since it is the total number in the adsorption volume, both bound and free The amount determined experimentally is the difference between the total amount of gas in the adsorbed phase and the amount in an equal volume of the gas phase At values of ϕ of the order of those considered in this paper and at small values of x, the difference between the total amount in the adsorbed phase and the experimental value of x is negligible

It is also important to note that the integration employed to give the adsorption isotherm assumes that the adsorbed and gaseous phases have the same dynamical state The adsorption isotherm will therefore be applicable to systems in which the adsorbed phase is similar to the model discussed above, but it is not applicable in the simple form, as we shall see later, to a system in which the adsorbed gas becomes a part of the solid phase and behaves, for example, as a Planck oscillator

The simple adsorption isotherm, given by equation (1), gives the amount adsorbed as proportional to the pressure, a result which agrees with experiment if x is small The deviations from this equation for larger values of x are quite considerable The Langmuir adsorption isotherm was one of the first successful attempts to account for these deviations It is important, however, to notice that the Langmuir adsorption isotherm can also be obtained if it is assumed that the adsorbing surface is uniform, can adsorb only one molecular layer, and that the adsorbed molecules are infinitely hard spheres which do not attract each other To this extent, therefore, his discussion is equivalent to the assumption that the intermolecular energy of interaction is $E(r) = +\infty$ $(r < \sigma)$ and $E(r) = 0$ $(r > \sigma)$, where σ is the molecular diameter and may be considered to take account of the imperfection of the adsorbed gas layer A theoretical derivation of an adsorption isotherm which takes account of both types of intermolecular forces is given below

The approach to this problem, using the Boltzmann equation itself, is difficult and cumbersome Fowler (1936*a*, p 253) points out that the theoretical treatment of imperfect gases in external fields of force is best accomplished by applying the laws of statistical mechanics only to elements of the gas and to supplement these by general thermodynamic theorems Following this line of attack we write the free energy for the gaseous and adsorbed phases as F_g and F_a respectively (cf Fowler 1936*a*, p 240)

Then
$$F_g = -kTN_g\left[\log b^g_{N_g}(T) + \log\frac{f_g(T)}{N_g} + 1\right], \tag{2}$$

$$F_a = -kTN_a\left[\log b^a_{N_a}(T) + \log\frac{f_a(T)}{N_a} + 1\right] \tag{3}$$

In these equations N_g and N_a are the total number of molecules in the gaseous and adsorbed phases respectively, and $f_g(T)$ and $f_a(T)$ are partition functions for molecular internal and translational energy, $b^g_{N_g}(T)$ and $b^a_{N_a}(T)$ are the corresponding partition function per molecule for the potential energy of the whole gas The b's are related to the strict partition function for the potential energy of the whole gas by $\{b(T)\}^N = B(T)$ The condition for equilibrium is that

$$(\partial F_a/\partial N_a)_{T,V} = (\partial F_g/\partial N_g)_{T,V}$$

If we assume the gaseous phase to be perfect, then $b^g_{N_g}(T) = V_g$, and the condition for equilibrium reduces to

$$\log V_g + \log\frac{f_g(T)}{N_g} = \log b^a_{N_a}(T) + \log\frac{f_a(T)}{N_a} + N_a\frac{\partial}{\partial N_a}\log b^a_{N_a}(T)$$

The partition function $b^a_{N_a}(T)$ can be evaluated in the following way If we assume that the adsorption potential (ϕ cal /mol) is uniform over the whole of the adsorbed phase then, neglecting the imperfection of the adsorbed layer,

$$b^a_{N_a}(T) = V_a e^{\phi/RT}$$

(cf Fowler 1936*a*, p 57)

The imperfection of the adsorbed layer can be taken into account by writing

$$b^a_{N_a}(T) = V_a\left(1 + \beta\frac{N_a}{V_a} + \gamma\frac{N_a^2}{V_a^2} + \quad\right)e^{\phi/RT},$$

and it will be shown later that the coefficients β, γ, etc are very simply related to virial coefficients Writing

$$b^a_{N_a}(T) = V_a B_{N_a}(T)\, e^{\phi/RT}, \tag{4}$$

we have

$$\frac{V_g f_g(T)}{N_g} = \frac{V_a B_{N_a}(T) f_a(T)}{N_a}\exp\left[N_a\frac{\partial}{\partial N_a}\{\log B_{N_a}(T)\} + \phi/RT\right],$$

$$\frac{C_g}{C_a} = \frac{f_g(T)}{f_a(T)\, B_{N_a}(T)}\exp\left[-\phi/RT - N_a\frac{\partial}{\partial N_a}\log B_{N_a}(T)\right]$$

The general equation assuming imperfection of the gaseous phase is

$$\frac{C_g}{C_a} = \frac{f_g(T)\,B_{N_g}(T)}{f_a(T)\,B_{N_a}(T)} \exp\left[-\phi/RT - N_a \frac{\partial}{\partial N_a} \log B_{N_a}(T) + N_g \frac{\partial}{\partial N_g} \log B_{N_g}(T) \right] \tag{5}$$

Equation (1) can be deduced readily from this if we use the same model of the adsorbed layer Since both gaseous and adsorbed layers are perfect we have

$$B_{N_g}(T) = B_{N_a}(T) = 1,$$

$$\frac{f_g(T)}{f_a(T)} = \frac{w_1}{w_2},$$

and therefore as before

$$x = p\frac{\alpha d}{RT}\frac{w_2}{w_1}e^{\phi/RT}$$

The process outlined above for evaluating equation (5) is typical of the general method, in which some definite model is chosen for the adsorbed phase and the partition functions corresponding to it evaluated In the further discussion given in this paper three models of the adsorbed phase will be considered In the first two the adsorbed gas will be assumed to behave as a highly compressed gas Proceeding from this, two special cases will be considered In the first model the adsorbed gas will be assumed to possess three translatory degrees of freedom This is equivalent to the assumption that the amplitude of any vibration of the adsorbed molecules perpendicular to the surface is much larger than the molecular diameter In the second model, the adsorbed gas will be assumed to be a two-dimensional gas, in which the molecules moving across the solid surface execute a vibratory motion perpendicular to the surface whose amplitude is of the order of the molecular diameter

In the third model the adsorbed molecules will be assumed to have no lateral but only vibratory movement As Lennard-Jones (1932) has shown, the solid surface is a region of varying adsorption potential even on a smooth surface, owing to the arrangement of atoms in a lattice structure (cf Fowler 1932) As the temperature is raised and the adsorbed molecules obtain sufficient energy to surmount the potential hills in the surface, this model passes over into the second model

CASE I

If the adsorbed phase is a three-dimensional gas, we can write

$$f_g(T)/f_a(T) = w_1/w_2$$

Assuming the gaseous phase to be perfect, we have

$$B_{N_g}(T) = 1$$

Equation (5) reduces then to

$$\frac{C_g}{C_a} = \frac{w_1}{w_2 B_{N_a}(T)} \exp\left[-\phi/RT - N_a \frac{\partial}{\partial N_a} \log B_{N_a}(T) \right]$$

Substituting $$B_{N_a}(T) = \left(1 + \beta \frac{N_a}{V_a} + \gamma \frac{N_a^2}{V_a^2} + \quad \right),$$

we get $$\log \frac{C_g}{C_a} = \log \frac{w_1}{w_2} - \frac{\phi}{RT} - \frac{2\beta N_a}{V_a} - \tfrac{3}{2}(2\gamma - \beta^2)\frac{N_a^2}{V_a^2}$$

If x is the number of g mol adsorbed per sq cm , d the thickness of the adsorbed layer and α the fraction of the surface available for adsorption, then

$$\log \frac{x}{p} = \log \frac{w_2}{w_1} \frac{\alpha d}{RT} + \frac{\phi}{RT} + \frac{2\beta N_0}{\alpha d} x + \frac{3}{2} \frac{N_0^2(2\gamma - \beta^2)}{(\alpha d)^2} x^2,$$

where ϕ is the adsorption potential per g mol and N_0 is Avogadro's number

To put this equation into a form suitable for application to experimental data it is necessary to determine the relation between β and γ and the virial coefficients for the gaseous phase In order to do this it will be assumed that the virial coefficients are the same for the adsorbed and gaseous phases, although later this will be modified to correct for the transition of molecules from a three-dimensional to a two-dimensional gas London (1931) shows that dispersion forces between molecules are to a first approximation additive, and in those gas-solid interfaces at which the van der Waals' forces are predominantly of the dispersion type it would appear to be reasonable therefore to assume that the dispersion forces are not very different in the adsorbed from those in the gaseous phase Such systems are those of argon, nitrogen and oxygen on platinum, which are considered in the next paper

Now $$p = -\left(\frac{\partial F_a}{\partial V}\right)_{T\ N_a} = -kN_a \frac{\partial}{\partial V}\left[\log \frac{b^a_{N_a}}{N_a} + \log f_a(T) + 1\right]$$

$$= -\frac{kN_a}{V_a}\left(1 - \frac{\beta N_a}{V_a} - \frac{(2\gamma - \beta^2) N_a^2}{V_a^2} - \quad \right)$$

$$= -\frac{Rx}{V_a}\left(1 - \frac{\beta N_0 x}{V_a} - \frac{(2\gamma - \beta^2) N_0^2 x^2}{V_a^2} - \quad \right)$$

But the virial coefficients B, C, may be defined by

$$p = \frac{RT\,x}{V_a}\left(1+\frac{Bx}{V_a}+\frac{Cx^2}{V_a^2}+\quad\right)$$

$$\beta = -B/N_0,$$

$$(2\gamma-\beta^2) = -C/N_0^2$$

$$\log\frac{x}{p} = \log\frac{w_2}{w_1}\frac{\alpha d}{RT}+\frac{\phi}{RT}-\frac{2Bx}{\alpha d}-\frac{3}{2}\frac{Cx^2}{(\alpha d)^2} \tag{6}$$

This equation can be derived equally simply by purely thermodynamic arguments, without using the notation of partition functions Bradley (1931) has developed an equation of the same form from such thermodynamic arguments Williams (1918) has also deduced an equation of the same type from a mixture of kinetic and thermodynamic arguments He proves that

$$\log\frac{x}{p} = A_0 - A_1 x,$$

where A_0 and A_1 are constants

It has been tacitly assumed that the adsorption potential ϕ is constant throughout the adsorption volume This assumption will be justifiable if monomolecular layers exist, when $(-\phi)$ will be the potential energy of the adsorbed molecule at its mean distance from the surface The simplest way of removing this restriction is to sum the x's for each value of ϕ which corresponds to each successive adsorption layer Neglecting the weight factors, for the first layer we have

$$\log\frac{x_1}{p} = \log\frac{\alpha_1 d_1}{RT}+\frac{\phi_1}{RT}-\frac{2B}{\alpha_1 d_1}x_1-\frac{3C}{2(\alpha_1 d_1)^2}x_1^2$$

For the second

$$\log\frac{x_2}{p} = \log\frac{\alpha_2 d_2}{RT}+\frac{\phi_2}{RT}-\frac{2B}{\alpha_2 d_2}x_2-\frac{3C}{2(\alpha_2 d_2)^2}x_2^2$$

Unfortunately, it is not possible to solve these equations in a simple manner for $(x_1+x_2+\quad)$ as a function of p, and some device of the type employed by Whipp (1933) for two layer Langmuir adsorption isotherms will be necessary before they can be applied to experimental results

CASE II ADSORBED PHASE TWO-DIMENSIONAL LATERAL FREE MOTION TOGETHER WITH VIBRATION PERPENDICULAR TO SOLID SURFACE

In this case it is no longer possible to assume that

$$f_g(T)/f_a(T) = w_1/w_2$$

We can write instead $\quad f_a(T) = f'_a(T) f_{\text{vib}}(T),$

where $f_{\text{vib}}(T)$ is the partition function for the vibrational energy of the adsorbed molecule in a direction perpendicular to the solid surface, and $f'_a(T)$ is the partition function for the two-dimensional translatory and rotatory energies The partition function for rotation has not been abstracted because on this model the rotational energy in the gaseous and adsorbed phase should be equal (cf Lennard-Jones 1932, p 340)

Now $$f_g(T) = \left(\frac{2\pi m kT}{h^3}\right)^{\frac{3}{2}}$$

For a two-dimensional gas

$$f'_a(T) = \frac{2\pi mkT}{h^2}$$

Assuming the vibrations are simple harmonic

$$f_{\text{vib}}(T) = \frac{(1 - e^{-ph\nu/kT})}{(1 - e^{-h\nu/kT})}$$

(Fowler 1936*a*, p 90) Here $ph\nu$ is the maximum vibratory energy the molecule can have without leaving the surface Since adsorption potentials are usually at least 5000 cal /g mol $e^{-ph\nu/kT}$ can be neglected and we have

$$f_{\text{vib}} = (1 - e^{-h\nu/kT})^{-1}$$

As before $B_{N_g}(T) = 1$, but $B_{N_a}(T)$ is now given by

$$B_{N_a}(T) = \left(1 + \beta'\frac{N_a}{A} + \gamma'\frac{N_a^2}{A^2} + \quad\right),$$

where β' and γ' are the values of β and γ for the two-dimensional gas and A is the area occupied by one gram molecule of gas Using the method of the Clausius virial the general equation of state of a two-dimensional imperfect gas can be shown to be (Mitchell 1935)

$$pA = NkT\left(1 - \frac{N^2\pi}{A}\int_0^\infty r(e^{-E(r)/kT} - 1)\,dr\right)$$

Assuming that $E = +\infty$ $(r < \sigma)$ and $E = 0$ $(r > \sigma)$, where σ is the radius of the molecule, we obtain

$$pA = NkT + \frac{1}{2}\frac{N^2\pi\sigma^2 kT}{A}$$

The corresponding equation for the three-dimensional gas is

$$pV = NkT + \frac{2}{3}\frac{N^2\pi kT\sigma^3}{V}$$

From this we see that the equation connecting B and B', where B' is the second virial coefficient for a two-dimensional gas, is approximately

$$B = \tfrac{4}{3}\sigma B'$$

Similarly we can write $\qquad C = \xi\sigma^2 C'$,

where ξ is a constant

Substituting the above derived values for the various partition functions in equation (5), we have

$$\frac{C_g}{C_a} = (2\pi m kT)^{\frac{1}{2}}\,(1 - e^{-h\nu/kT})\,e^{-\phi/RT + \frac{2B}{A} + \frac{3C'^2}{2A^2}}$$

and

$$\log\frac{x}{p} = \log\frac{\alpha\sigma}{RT}\frac{w_2}{w_1} - \log\frac{(2\pi m kT)^{\frac{1}{2}}}{h} - \log(1 - e^{-h\nu/kT}) + \frac{\phi}{RT} - \frac{3}{2}\frac{B}{\alpha\sigma}x - \frac{\xi C}{2(\alpha\sigma)^2}x^2 \qquad (7)$$

This equation is very similar in type to equation (6) developed from the simpler model

CASE III ADSORBED GAS A PLANCK OSCILLATOR OF 3 DEGREES OF FREEDOM

Here also it is not possible to write

$$f_g(T)/f_a(T) = w_1/w_2$$

The adsorbed molecule can be assumed to be a Planck oscillator with three degrees of freedom, and we get

$$f_a(T) = f_{\text{vib}}(T) = (1 - e^{-h\nu/kT})^{-3},$$

$$f_g(T) = (2\pi m kT)^{\frac{3}{2}}/h^3$$

As before we can write

$$B_{N_a}(T) = \left(1 + \frac{\beta''}{A} + \frac{\gamma''}{2A^2} + \quad\right)$$

Substituting in (5) we obtain

$$\frac{C_g}{C_a} = \frac{w_1}{w_2}\frac{(2\pi mkT)^{\frac{3}{2}}}{h^3}(1-e^{-h\nu/kT})^3 e^{-\phi/RT+\frac{B''}{A}+\frac{C''}{2A^2}+} \quad ,$$

$$\log\frac{x}{p} = \log\frac{\alpha\sigma}{RT}\frac{w_2}{w_1} - \log\frac{(2\pi mkT)^{\frac{3}{2}}(1-e^{-h\nu/kT})^3}{h^3} + \frac{\phi}{RT} - \frac{2B''}{A} - \frac{3C''}{2A^2}, \qquad (8)$$

where B'' and C'' are constants similar in type to virial coefficients

It is interesting at this stage to derive the simplified versions of equations (7) and (8) which can be applied if the adsorbed gas phase is perfect, and compare the results with equation (1) From (7) we get

$$x = \frac{\alpha d}{RT}\frac{w_2}{w_1}\frac{(2\pi mkT)^{\frac{1}{2}}(1-e^{-h\nu/kT})}{h}e^{\phi/RT}p,$$

and from (8) $$x = \frac{\alpha d}{RT}\frac{w_2}{w_1}\frac{(2\pi mkT)^{\frac{3}{2}}}{h^3}(1-e^{-h\nu/kT})^3 e^{\phi/RT}p$$

This makes quite evident the errors which are obtained if the Boltzmann equation is not applied with sufficient discrimination

SHAPE OF THE THEORETICAL ADSORPTION ISOTHERMS

The shapes of the p/x isotherms given by equations (6), (7) and (8) are obviously similar It is easy to see that at small values of x, x is proportional to p As x increases the progress of the curve depends on the sign of B, the second virial coefficient If B is positive, which corresponds with that portion of the pV/p curve for gases at which $\partial(pV)/\partial p$ is negative, then as x increases the adsorption isotherm becomes convex to the pressure axis At this stage the nett intermolecular forces are attractive If B is negative, then the isotherm is concave to the pressure axis and the nett intermolecular force in the adsorbed layer is repulsive At high values of x, the adsorbed gas will always be at that part of the pV/p curve where $\partial(pV)/\partial p$ is positive

It should, perhaps, be pointed out that none of the isotherms, either (6), (7) or even (8) which approaches most nearly to Langmuir's model, gives a saturation maximum, because the adsorbed molecules are not assumed to be incompressible For all practical purposes, of course, the isotherms give a saturation maximum owing to the relatively slight compressibility of molecules

THE DIFFERENTIAL HEAT OF ADSORPTION AND THE ADSORPTION POTENTIAL

The most commonly determined heat of adsorption is the integral heat of adsorption Q The differential heat is then defined as $(\partial Q/\partial x)$, which we will write as q Comparatively few direct measurements of adsorption heats have been made and the normal method of procedure is to apply the equation

$$(\partial \log p/\partial T)_x = -q_x/RT^2 \tag{9}$$

to two different $\phi_1 T$ points on the same x isostere The above equation is usually considered to be a corollary to the Clausius-Clapeyron equation Coolidge (1926), after a critical discussion, concludes that this equation is satisfactory McBain (1932), however, draws attention to the difficulty that where saturation maxima vary with temperature it is impossible to construct the isostere when x approximates to x_s, the saturation value and the equation appears to break down He suggests that this is due to the strains set up in the solid during adsorption and that the extent of this strain varies with temperature Further difficulties are obtained when differential heats so calculated are compared with directly determined values Coolidge found they agreed only to a first approximation, the calculated values being somewhat less than the experimental A similar discrepancy has been reported by Pearce and Taylor (1931)

The above difficulties are entirely removed if the equation is applied as it should be to the concentration $(x/\alpha d)$ isosteres instead of x isosteres, for (αd) is not independent of temperature, in fact, as will be shown in a later paper,

$$\alpha d = (\alpha d)_0 \, e^{-l/\beta T} \tag{10}$$

It is not necessary to derive the exact equation thermodynamically because it can be obtained without trouble from the general adsorption isotherm (equation (5)) In this equation the $B(T)$'s determine the potential energy of a molecule due to molecular interaction and we can write

$$-\frac{W_g}{RT} = \log B_{N_g}(T) + N_g \frac{\partial}{\partial N_g} \log B_{N_g}(T),$$

$$-\frac{W_a}{RT} = \log B_{N_a}(T) + N_a \frac{\partial}{\partial N_a} \log B_{N_a}(T),$$

where W_g and W_a are the potential energies with reference to a perfect gas as the standard state and are functions of C_g and C_a respectively Then

$$\log\frac{x}{p} = \log\frac{x}{\alpha d} + \frac{5}{2}\log T + \log\frac{(2\pi m)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3} - \log\frac{(2\pi m)^{s'/2}}{h^{s'}}$$

$$+\log(1-e^{-h\nu/kT})^{s''} - \frac{\phi + W_g - W_a}{RT},$$

where s' and s'' are the appropriate degrees of freedom for translation and vibration in the adsorbed phase

Integration of the Clausius-Clapeyron equation gives

$$\log\frac{x}{p} = \log\frac{x}{\alpha d} + \frac{(C_l)_g - (C_l)_a}{R}\log T - \int_0^T \frac{dT}{RT^2}\int_0^T C_{\text{vib}}\, dT + \frac{q_0 - W_g + W_a}{RT} + \iota \tag{11}$$

Here $(C_l)_g$ and $(C_l)_a$ are the constant part of the specific heat for gaseous and adsorbed phases, C_{vib} is the specific heat associated with the vibration of the adsorbed molecule and ι is the chemical constant The specific heats due to rotation and internal vibrations of the molecules have been assumed to be equal in both phases Comparing (9) and (11)

$$-\phi + W_g + W_a = q_0 - W_g + W_a$$

and

$$\iota = \log\frac{(2\pi m)^{\frac{3-s}{2}} k^{\frac{5-s}{2}}}{h^{3-s}}$$

q_0 is here the heat of adsorption when both gaseous and adsorbed phases are perfect and at absolute zero, and

$$q = (q_0 - W_g + W_a) - \int_0^T C_{\text{vib}}\, dT - \{(C_l)_s - (C_a)_s\}\, T, \tag{12}$$

where q is the differential heat of adsorption and is given by

$$(\partial \log p/\partial T)_{(x/\alpha d)} = -q/RT^2$$

This gives us the relation, therefore, between ϕ and q We can rewrite (12) as

$$q = q_T - W_g + W_a, \tag{13}$$

where q_T is the differential heat of adsorption at temperature T under conditions such that W_g and W_a are zero, i e low concentrations in both adsorbed and gaseous phases

The error which arises where x isosteres are used to calculate q instead of $(x/\alpha d)$ isosteres can be demonstrated readily From (11), (12) and (13)

$$\left(\frac{\partial \log p}{\partial T}\right)_x = \frac{\partial}{\partial T}\left(\log \frac{1}{\alpha d}\right)_x - \frac{q}{RT^2}$$

From (10) we have $\alpha d = (\alpha d)_0 \, e^{-\beta/T}$,

where β is a positive quantity, so that

$$\frac{\partial}{\partial T}\left(\log \frac{1}{\alpha d}\right) = +\frac{\beta}{RT^2}, \quad q_x = q - \beta$$

Values calculated from the x isosteres are therefore less than the true values by the amount β q_x is therefore greater than q, in agreement with the experimental results of Pearce and Taylor

THE DECREASE IN q WITH INCREASE OF x

As a general rule, the differential heat of adsorption decreases with increase in amount adsorbed The theoretical relation between x and q can be readily derived From (9), (11) and (12), we have

$$\log x/p = \log \alpha d + Tf'(T) + f''(T) - q/RT,$$

where

$$f'(T) = \left\{\frac{5}{2}T + \frac{se^{-h\nu/kT}}{(e^{-h\nu/kT}-1)}\frac{h\nu}{kT^2}\right\},$$

$$f''(T) = \left[\frac{5}{2}\log T + \log\frac{(2\pi m)^{\frac{3}{2}}\,k^{\frac{3}{2}}}{h^3} - \log\frac{(2\pi m)^{s/2}}{h^s} + \log(1-e^{-h\nu/kT})^s \right]$$

There are no satisfactory data existing which enable a test to be made of this equation, for no one has measured heats of adsorption on plane surfaces The calculated values reported in the literature are erroneous for the reasons mentioned above The direct relation between q and x is given by (13) If as usual $W_g = 0$,

$$q = q_T - RT\left(\frac{\beta x}{\alpha d} + 2\gamma\frac{x^2}{(\alpha d)^2} + \quad\right)$$

If this decrease is due to molecular interaction and not to heterogeneity of the surface field β must be positive Therefore B, the second virial coefficient, must be negative, indicating a nett repulsive force between molecules in the adsorbed layer

SUMMARY

The generalized adsorption isotherm is shown to be

$$C_g/C_a = \frac{f_g(T)\,B_{N_g}(T)}{f_a(T)\,B_{N_a}(T)} \exp\left[-\phi/RT + N_g \frac{\partial}{\partial_{N_g}} \log B_{N_g}(T) - N_a \frac{\partial}{\partial_{N_a}} \log B_{N_a}(T)\right],$$

where C_a and C_g are the molecular concentrations in the adsorbed and gaseous phases respectively, ϕ is the adsorption potential and $f_g(T)$, $f_a(T)$, $B_{N_g}(T)$ and $B_{N_a}(T)$ are partition functions These functions have been evaluated for three special cases in which the adsorbed phase is (*a*) a three-dimensional gas, (*b*) a gas of two-dimensional lateral mobility whose molecules vibrate in a plane perpendicular to the plane of lateral mobility, and (*c*) a group of Planck oscillators with three degrees of freedom

The relation between ϕ and q, the differential heat of adsorption, is deduced and it is shown that the common practice of calculating from the equation

$$(\partial \log p/\partial T)_x = -q_x/RT^2$$

is erroneous The correct equation is

$$(\partial \log p/\partial T)_{(x/\alpha d)} = -q/RT^2,$$

where $(x/\alpha d)$ is the concentration in the adsorbed layer

My thanks are due to Professor R H Fowler for much kind advice and assistance during the preparation of this paper

REFERENCES

Bradley, R S 1931 *Phil Mag* Ser 7, **11**, 690–6
Coolidge, A S 1926 *J Amer Chem Soc* **48**, 1795–814
Fowler, R H 1932 *Trans Faraday Soc* **28**, 419–20
— 1936*a* "Statistical Mechanics" Cambridge, 2nd ed
— 1936*b* *Proc Camb Phil Soc* **32**, 144–51
Kruyt and Moddermann 1930 *Chem Rev* **7**, 329
London, F 1931 *Z phys Chem* B **11**, 222–51
Lennard Jones, J E 1932 *Trans Faraday Soc* **28**, 333–59
McBain, J W 1932 "Sorption of Gases", p 141
Mitchell, J S 1935 *Trans Faraday Soc* **31**, 980–6
Pearce, J N and Taylor, A L 1931 *J Phys Chem* **35**, 1091–103
Peierls, R 1936 *Proc Camb Phil Soc* **32**, 471–6
Whipp, B 1933 *Proc Roy Soc* A, **141**, 216–32
Williams, A M 1918 *Proc Roy Soc*, *Edin*, **38** 23–9

The adsorption of argon, nitrogen and oxygen on smooth platinum foil at low temperatures and pressures

By F J Wilkins

Research Department, Imperial Chemical Industries, Billingham

(*Communicated by R H Fowler, F R S —Received* 20 *September* 1937)

The work to be described in this paper is, in effect, an extension of the earlier researches of Langmuir (1918), who measured the adsorption of argon, nitrogen, oxygen, carbon monoxide, carbon dioxide and methane on mica and glass at pressures up to 0 1 mm and temperatures of 90° and 155° K He also investigated the adsorption of these gases on a smooth platinum foil surface, but except for a slow irreversible adsorption of oxygen at temperatures above 0° C, he was unable to detect any adsorption

In order to interpret his experimental results, Langmuir deduced his well-known adsorption isotherm

$$x = abp/(1+ap),$$

where x is the amount adsorbed at a pressure p and temperature T Of the constants b is equal to the saturation maximum, a is equal to $a_0 e^{\phi/RT}$ where a_0 is a constant and ϕ is the adsorption potential

Langmuir showed that this equation was satisfied by his experimental results It has since received widespread application One of the outstanding successes was obtained in the careful work of McBain and Britton (1930) on the sorption of the vapours of nitrous oxide and ethylene on charcoal

Failure of the values of x and p to satisfy this equation is usually attributed to the composite nature of the adsorbing surface There are, however, two serious difficulties which cannot be disposed of so easily Both concern the constant b, the saturation maximum If Langmuir's equation is correct, the value of b should be the amount of gas required to cover the adsorbing surface with a layer one molecule thick The following table gives Langmuir's and Bawn's figures for the fraction of the adsorbing surface covered when various gases were adsorbed on mica, Huckel (1928) Bawn (1932)

With the exception of Bawn's figures for carbon monoxide the surface is only partially covered at saturation An even more serious criticism has been brought forward by Huckel (1928), who pointed out that theoretically b should be independent of temperature Actually, as the above table shows, b is very markedly reduced by increase of temperature Zeise (1928) showed empirically that $b = b_0(1 - kT)$, where b_0 and k are constants

TABLE I

	Langmuir		Bawn		
Gas	$T = 90°$ K	$T = 155°$ K	$T = 90°$ K	193° K	240° K
N_2	0 22–0 26	0 08	—	—	—
CH_4	0 64–0 86	0 17	—	—	—
CO	0 38	0 13	1 13	0 82	—
A	0 17	0 06	0 30	—	—
O_2	0 11	0 03	0 23	0 16	—
CO_2	—	0 41–0 47	—	—	—
Acetone	—	—	0 32	—	0 33

In order to investigate these problems further it was decided to carry out measurements similar to those made by Langmuir Instead of using mica and glass it was decided to use platinum foil as the adsorbing surface Mica and glass, although readily obtainable in thin sheets with smooth surfaces, have the drawback that they possess so complicated a lattice structure that the arrangement of the surface atoms is not accurately known Platinum, on the other hand, is free from this objection and, moreover, can be obtained as a smooth foil which can be cleaned vigorously without impairing the surface smoothness As it was almost certain that Langmuir s failure to measure adsorption on platinum at low pressure and temperatures was due only to the insensitivity of his apparatus, it was decided to use platinum as an adsorbent After a few trial experiments, an apparatus was designed capable of measuring the adsorption of argon, nitrogen and oxygen on platinum at temperatures between 90° and 190° K

EXPERIMENTAL

The arrangement of the apparatus is shown in fig 1 The adsorption vessel A is made of pyrex and is connected via a pyrex-glass seal to a Pirani gauge from which it can be isolated by means of a tap The Pirani gauge in its turn is connected to a vacuum line, a McLeod gauge and a storage vessel which contained the purified gas required for the experiments All glass or pyrex used for the apparatus was carefully cleaned with hot chromic acid and distilled water The vacuum for the adsorption section was obtained from an Apiezon oil diffusion pump backed by a Hyvac These were separated from the tap isolating the Pirani gauge system by a trap, surrounded by liquid air day and night, and a P_2O_5 tube The vacuum for the McLeod gauge section, which was used during the calibration of the Pirani gauge, was obtained from a two-stage Volmer mercury diffusion pump set, backed with a Hyvac

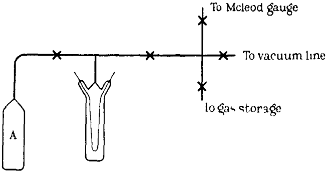


FIG 1

Great care was taken to keep the apparatus free from mercury and grease Mercury vapour from the McLeod gauge was condensed out in a liquid air trap and the traces that passed through the trap entered the Pirani gauge chamber only, as all calibrations were made before the adsorption vessel was sealed on All taps were carefully ground so that mercury sealing was unnecessary and they were greased with Apiezon low vapour pressure grease

PIRANI GAUGE

The filament of the Pirani gauge was made of flattened nickel wire (Ellet and Zabel 1931) It was operated according to the method of Campbell This method gives at low pressures a linear relation between the pressure p and the square of the voltage across the bridge

$$p = k(V^2 - V_0^2)/V_0^2,$$

where V is the bridge voltage at pressure p and V_0 the voltage *in vacuo* Calibrations were carried out over the range 5×10^{-5} to 10^{-1} mm At the higher pressures used in these experiments the above equation departed from linearity and calibrations were made at frequent pressure intervals The voltmeter used was a Crompton universal indicator

VOLUME MEASUREMENTS

Volume measurements were made with reference to a bulb whose volume had been measured by filling it with water, using hydrogen as the reference gas Before making them, the whole of the apparatus was very carefully baked out

TEMPERATURE MEASUREMENT

Experiments were carried out at various temperatures between 80° and 195° K These were obtained by using constant temperature baths of liquid nitrogen, liquid air, liquid oxygen, liquid methane, liquid ethylene and solid carbon dioxide and acetone The temperatures of these baths were measured with a copper constantan thermocouple, which was calibrated frequently against a platinum resistance thermometer, which had been standardized at the National Physical Laboratory The temperature measurements were self-consistent to within 0 5°

PREPARATION AND PURIFICATION OF GASES

(a) *Oxygen*

The oxygen was prepared by heating A R potassium permanganate It was purified by treatment with caustic potash and then by storage in a vessel surrounded with liquid air

(b) *Nitrogen*

The nitrogen was kindly supplied by Dr E J B Willey, who reported that the sample was so pure that it refused to give the active nitrogen afterglow It was prepared from air and therefore contained the usual traces of inert gases

(c) *Argon*

The argon was obtained from a cylinder supplied commercially It was purified using the method of Baxter and Starkweather, which consists in fractionating it from chabazite between the temperatures of 90° and 190° K This method is reported to give a gas which is spectroscopically pure

(d) *Hydrogen*

The small quantities of hydrogen needed were made as they were required by heating with a hydrogen flame a portion of a hollow palladium tube which was sealed into the apparatus

KNUDSEN THERMAL TRANSPIRATION CORRECTION

Over the range of pressure used in these experiments, 5×10^{-5} to 10^{-1} mm , the thermal transpiration correction is very important It was impossible to reduce it to negligible proportions by working with large diameter tubes, as the volume of the apparatus was reduced to a minimum in order to achieve the maximum sensitivity It was also impossible to work over the square root range, as this would have required a capillary tube connection between the adsorption vessel and the Pirani gauge, which would have made the task of baking out the adsorption vessel very difficult indeed The tube connecting the adsorption vessel and Pirani gauge in the final series of experiments had an internal diameter of 4 mm The Knudsen correction was determined for argon, nitrogen and oxygen at all temperatures used and over the whole pressure range

Langmuir estimated his correction by determining it for hydrogen, which is not adsorbed, and then assuming it was the same for all other gases at those pressures at which the mean free paths were equal This assumption is true only over a narrow pressure range at the lowest pressures

THE MEASUREMENT OF THE THERMAL TRANSPIRATION CORRECTION

In order to measure the thermal transpiration correction, the apparatus was first carefully baked out and evacuated During this operation, the pyrex adsorption vessel was kept at a temperature of 500° C for 24 hours After isolating the pumps and closing the tap between the adsorption vessel and the Pirani gauge, gas was admitted to the gauge at a measured pressure, p_1 It was then expanded to the adsorption vessel, which contained no platinum and which was placed in a constant temperature bath The new pressure p_2 was measured after equilibrium had been reached From these two figures the apparent pressure, p_3, in the adsorption vessel was calculated assuming that there was no adsorption Then

$$\frac{p_1 V_1}{RT_1} = \frac{p_2 V_1}{RT_1} + \frac{p_2 V_2}{RT_1} + \frac{p_3 V_3}{RT_2},$$

where V_1 is the volume of the Pirani gauge, V_2 the volume of the connecting tube between the tap and the level of the liquid in the constant temperature bath, and V_3 the volume of the adsorption vessel immersed in the bath T_1 is the air temperature, T_2 the bath temperature, R the gas constant

Now if x is the mass of gas adsorbed on the walls of the adsorption vessel, at the apparent pressure of p_3 we have

$$\frac{p_3 V_3}{RT_2} = x + \frac{p_0 V_3}{RT_2},$$

where p_0 is the true pressure in the adsorption vessel. p_3 was measured for all gases over the whole range of temperature and pressure. The experiments were then repeated with the adsorption vessel filled with pyrex glass tubing of known area, which enabled a series of pressures p_4 to be determined

$$\frac{p_4 V_4}{RT_2} = \alpha x + \frac{p_0 V_3}{RT_2},$$

where V_4 is the volume of the adsorption vessel minus the volume of the added tubing and α is the ratio between the surface area of the adsorption vessel together with the tubing and the surface area of the adsorption vessel.

From these two series of experiments values of p_0 and x were determined for all values of p_2. In actual fact the values of x were so surprisingly small that they could be neglected except for temperatures of 90° K. or less. The Knudsen thermal transpiration correction $t = p_0/p_2$ was then calculated for all values of p_2.

In Table II are given the observed results for the transpiration correction for nitrogen and also the values calculated from the measured transpiration correction for hydrogen, assuming that the corrections are equal at the same mean free path. It is clear that Langmuir's assumption is not correct.

TABLE II

p_2 in mm 10^3	t measured	t calculated
0.5	0.54	—
1.0	0.61	0.63
1.5	0.68	0.64
2.0	0.73	0.66
2.5	0.77	0.68
3.0	0.79	0.69
3.5	0.81	0.71
4.0	0.82	0.72
5.0	0.84	0.75
6.0	0.86	0.78
7.0	0.87	0.80

ADSORPTION EXPERIMENTS

The platinum used in these experiments consists of a roll of smooth platinum foil, 9600 sq. cm. in area. This was carefully cleaned, first with hot nitric acid and then with hot chromic acid It was finally washed thoroughly with distilled water and placed in the pyrex adsorption bulb All the water adhering to the foil, with the exception of the last traces, was removed by heating under vacuum in a separate apparatus The adsorption bulb was then sealed into its final position in the apparatus and baked out at a temperature of 500° C until, after isolating the pumps, no measurable pressure developed over a period of 12 hours This baking-out process took several days

The method of carrying out the adsorption experiments was almost identical with that used to determine the thermal transpiration coefficients The apparatus was baked out for 12 hours at 500° C Gas at a known pressure was expanded from the Pirani gauge to the adsorption vessel, which was then isolated and more gas was added to the Pirani gauge This in its turn was allowed to expand into the adsorption vessel In this way a series of points on each isotherm was determined At low temperatures the amounts adsorbed were so large that, even when the Pirani gauge was filled with gas at as high a pressure as the calibration curve allowed (10^{-1} mm), on expansion to the adsorption vessel, the equilibrium pressure was too small for the gauge to measure Under these conditions the error of the measurement is equal to the error in measuring the pressure, which is $\pm 2\%$ At the high temperatures 170° and 190° K , when the amounts adsorbed were less, the errors became larger

RESULTS

The experimental results are given in Tables III–XIV The curves for nitrogen are plotted in fig 2 to illustrate the smoothness of the experimental curves and the change in the shape of the curve with change of temperature

Unfortunately, owing to a serious breakdown in the apparatus and to the short time available for completing this investigation it was not possible to carry out more than one check experiment If the results given in Tables III and IV for the adsorption of nitrogen at 83 5° K are compared it is found that the values of x given in Table III are about 4 % lower than those given in Table IV This agreement can be considered to be satisfactory in view of the nature of the experiments and the possible error in temperature measurement of 0 5° K Further, the self-consistency of the experimental results as judged by the closeness with which they can be placed along a smooth curve is also very satisfactory Only in the case of the experiments at 170° and 193° K , where the amounts adsorbed are very small, do the errors become at all serious

NITROGEN EXPERIMENTS

TABLE III

Bath Liquid air
Temperature = 83 5° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{10}
0 813	1 99
1 08	2 21
1 75	2 44
2 24	2 61

TABLE IV

Bath Liquid air
Temperature = 83 5° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{10}
0 069	1 19
0 325	1 75
0 533	2 00
0 996	2 27
1 52	2 51
2 13	2 69
2 88	2 89
3 71	3 06
4 37	3 19
5 14	3 34
6 01	3 47
6 93	3 59

TABLE V

Bath Liquid oxygen
Temperature = 90 5° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{10}
0 34	0 98
0 78	1 25
1 51	1 52
2 49	1 70
3 64	1 90
6 14	2 16
8 45	2 37
10 70	2 51
12 40	2 59

TABLE VI

Bath Liquid methane
Temperature = 112 2° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{11}
0 305	2 32
1 04	4 26
2 20	5 58
3 87	6 82
5 87	7 55
7 95	8 15
10 2	8 34
12 2	8 44
14 5	8 60

TABLE VII

Bath Liquid ethylene
Temperature = 170 4° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{12}
0 687	2 04
1 44	7 72
2 52	10 6
3 98	18 2
5 56	20 2
7 40	23 9
10 10	23 9
12 9	26 7

TABLE VIII

Bath Solid carbon dioxide acetone
Temperature = 194 5° K

p in mm 10^2	x in g mol /sq cm of Pt 10^{14}
0 877	2 34
2 04	3 18
3 17	6 72
4 55	5 55

ARGON EXPERIMENTS

TABLE IX

Bath Liquid nitrogen
Temperature = 77 3° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{10}
0 496	2 81
0 835	3 34
1 99	4 25
3 75	5 02
5 59	5 81
7 54	6 25
9 82	6 80
12 1	7 30
15 2	8 08

TABLE X

Bath Liquid oxygen
Temperature = 90 5° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{10}
1 40	1 33
2 74	1 75
4 27	1 99
6 11	2 25
7 97	2 45
10 14	2 71
12 3	2 92
15 3	3 15
18 8	3 30
21 5	3 43

TABLE XI

Bath Liquid methane
Temperature = 112 2° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{11}
0 80	3 33
1 95	5 59
3 15	6 97
5 50	8 73
8 97	10 8
12 1	12 2
14 0	13 1
16 1	14 0

TABLE XII

Bath Liquid ethylene
Temperature = 172 2° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{12}
1 01	2 00
2 16	7 57
3 91	10 40
6 66	17 7
10 5	19 7
14 5	23 2
18 8	26 0

OXYGEN EXPERIMENTS

TABLE XIII

Bath Liquid oxygen
Temperature = 90 5° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{10}
0 201	1 20
0 442	1 52
1 34	2 07
2 66	2 49
4 52	2 87
6 64	3 21
8 48	3 41
10 20	3 61
12 1	3 84

TABLE XIV

Bath Liquid ethylene
Temperature = 172° K

p in mm 10^3	x in g mol /sq cm of Pt 10^{11}
0 806	0 518
1 78	0 929
2 99	1 58
4 46	1 64
6 13	2 57
8 02	3 28
10 5	3 81

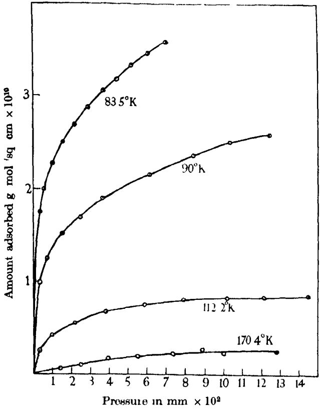


FIG 2 Nitrogen isotherms

DISCUSSION

Application of the Langmuir isotherm to the results

At the lower temperatures all three gases give adsorption isotherms which are not concave enough to the pressure axis to be typical of the Langmuir adsorption isotherm In each case there appears to be a middle part of the isotherm which is almost linear (see fig 2) The Langmuir equation has been applied to the results in the usual way by plotting p/x against p, which should give a straight line Such curves, except for some experiments at 112° and 170° K , are concave to the p-axis and indicate that more gas is adsorbed in the later higher pressure stages than is to be expected (see figs 3 and 4) The curves for nitrogen are particularly interesting because they show that the Langmuir equation is obeyed at 112° and 170° K That this is not due to a masking of the curvature caused by experimental inaccuracy is proved by the smoothness of the isotherm at 112° K and also by the way in which the individual points of the plot of p/x against p lie so accurately on a straight line The evidence for argon and oxygen is similar, but not so clear in that argon, for example, still shows a definite departure from the Langmuir isotherm at 112° K For all three gases it can be said that the higher the temperature at which adsorption occurs the more closely is the Langmuir equation obeyed The excess adsorption which occurs in all cases over and above that expected from the equation is to be attributed without doubt to intermolecular forces of attraction which will be more apparent at low temperatures

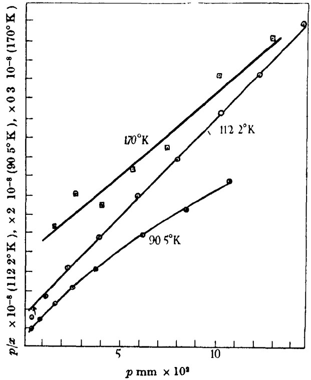


FIG 3 Nitrogen

The saturation maximum

The values of x_s, the saturation maximum, can of course be easily calculated for those experiments whose results satisfy the Langmuir equation For the other cases the following procedure has been adopted A straight line has been drawn through the last three points of the p/x, p curve In every case these points lay quite accurately on this line From the slope of this line the value of x_s is calculated in the usual way This method is probably not very greatly in error provided that the values of x corresponding to these three points are at least 50–60 % of the value of x_s, because at this stage in the adsorption the concentration in the adsorbed layer is so large that the forces of repulsion between molecules predominate in determining the adsorption equilibrium The results are given in Table XV and it is easy to see that x_s decreases rapidly with increase of temperature in agreement with the results of Bawn and Langmuir

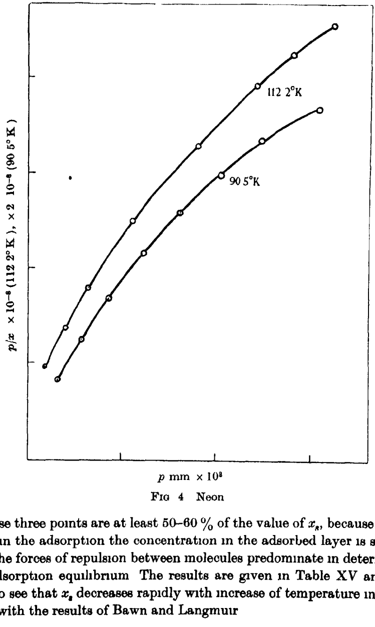


FIG 4 Neon

TABLE XV SATURATION MAXIMA

Gas T° K	Nitrogen x, g mol / cm $^2\times 10^{10}$	Oxygen x, g mol / cm $^2\times 10^{10}$	Argon x, g mol / cm $^2\times 10^{10}$
77 3	—	—	13 0
83 5	4 6	—	—
90 5	3 3	5 4	4 3
112 2	1 1	—	2 3
170 4	0 42	0 95	0 39

Application of the virial adsorption isotherm to the results

In an attempt to elucidate further the p/x relationships which have been found experimentally the virial adsorption isotherms which were derived in the previous paper have been applied to the experimental results (see equations (6), (7) and (8) of the previous paper)

In order to do this it was necessary to fit the experimental results to an equation of the type

$$\log x/p = k + bx + cx^2 +$$

As a first attempt Mr R R M Mallock substituted the experimental values in the above equation and solved directly for the coefficients on his calculating machine The range of values of x, which owing to the experimental difficulties was limited, was too small to enable any coefficients to be determined with any degree of exactness at all The coefficients were therefore determined in the following way For small values of x we have

$$\log x/p = k + bx$$

If this equation is applied to values of x and p outside the range over which this equation is valid, then

$$\log x/p = k_1 + b_1 x$$

If the values of x are not too large, then we can write

$$k_1 = k + cx$$

and $$b_1 = b_0 + dx$$

Substituting $$\log x/p = k + (c + b_0)\, x + dx^2$$

These equations were solved from three values of x and p taken from the smoothed curve over the range occupied by the first three x and p readings on the $\log \frac{x}{p}\Big/ x$ curve At the higher values of x, the above parabolic equation gave values of log x/p which were too low It was always possible to obtain a perfect fit over the whole range by adding a term in x^4 This method was of course employed only for the measurements at or below 90 5° K At the high temperatures the curve fitting was simplified (although the experimental results, particularly at 170° K , were not very accurate), because owing to the small amounts adsorbed, the lower values of x fitted the equation

$$\log x/p = k + bx$$

The higher values of x were easily dealt with by the addition of a term in x^2
The equations so calculated are given below

Argon

$T = \ 77\ 3^\circ\mathrm{K} \quad \log x/p = -13\ 5 - 1\ 40 \times 10^{10}x + 0\ 92 \times 10^{19}x^2 - 1\ 4 \times 10^{36}x^4$

$T = \ 90\ 5^\circ\mathrm{K} \quad \log x/p = -15\ 6 - 2\ 65 \times 10^{10}x + 0\ 44 \times 10^{20}x^2 - 2\ 8 \times 10^{38}x^4$

$T = 112\ 2^\circ\mathrm{K} \quad \log x/p = -18\ 6 - 2\ 05 \times 10^{10}x + 0\ 23 \times 10^{20}x^2$

$T = 170\ 4^\circ\mathrm{K} \quad \log x/p = -21\ 5 - 4\ 9\ \ \times 10^{10}x$

Nitrogen

$T = \ 83\ 5^\circ\mathrm{K} \quad \log x/p = -12\ 0 - 3\ 35 \times 10^{10}x + 0\ 39 \times 10^{20}x^2 - 0\ 92 \times 10^{38}x^4$

$T = \ 90\ 5^\circ\mathrm{K} \quad \log x/p = -14\ 0 - 4\ 43 \times 10^{10}x + 0\ 90 \times 10^{20}x^2 - 1\ 34 \times 10^{38}x^4$

$T = 112\ 2^\circ\mathrm{K} \quad \log x/p = -17\ 8 - 3\ 56 \times 10^{10}x - 4\ 61 \times 10^{18}x^2$

$T = 170\ 4^\circ\mathrm{K} \quad \log x/p = -21\ 0 - 4\ 6\ \ \times 10^{10}x$

Oxygen

$T = \ 90\ 5^\circ\mathrm{K} \quad \log x/p = -13\ 4 - 3\ 42 \times 10^{10}x + 0\ 60 \times 10^{20}x^2 - 0\ 85 \times 10^{38}x^4$

$T = 172\ 2^\circ\mathrm{K} \quad \log x/p = -21\ 0 - 2\ 76 \times 10^{10}x$

x is in g mol /sq cm and p in mm of mercury

Having determined values of k and b it is possible by substituting in equations (6), (7) and (8) of the previous paper to calculate (αd) or ($\alpha\sigma$) the volume of the adsorbed phase and, ϕ providing B, the virial coefficient for the gas phase, is known α is the fraction of surface area available for adsorption and d is the thickness of the adsorbed layer in one case, and in the other it is assumed to be equal to the molecular diameter of the adsorbed molecule σ (see previous paper)

Evaluation of the virial coefficients

The values of the second virial coefficient B can be calculated from gas data using Lennard-Jones' method (Fowler 1936, p 299) For both argon and nitrogen at all temperatures used in this investigation B is negative, i e the molecular attractive forces are greater than those of repulsion The experimental results however show that the forces of repulsion predominate (since dp/dx increases with increase of x) and it is evident that over the observed range of experiment virial coefficients of higher order than the second are the important factors The adsorbed gas is in fact on that part of the isotherm which is above the Boyle point and to which the equation $p(V-b) = kT$ applies approximately Since repulsive forces predominate it is possible to evaluate roughly a coefficient by the Lennard Jones method neglecting the attractive forces This is broadly what has been done in applying the virial coefficients to calculation of crystal parameters, except of course that the assumption that repulsive forces only are important is much more nearly true for the crystal state than for the adsorbed layer The values of B have therefore been calculated for argon and nitrogen assuming an $r^{-9\,3}$ law for the repulsive force No data exist which allow us to calculate B for oxygen

Having evaluated B it is possible to substitute in the adsorption isotherms (equations (6) and (7)) and evaluate (αd) or $(\alpha\sigma)$ For equation (8) the simplest assumption is to put $B' = B''$ This is not exactly true, but in the special case under consideration the error is perhaps not large The molecular diameters of argon, oxygen and nitrogen are $3\,64 \times 10^{-8}$, $3\,62 \times 10^{-8}$ and $3\,80 \times 10^{-8}$ cm respectively The platinum atoms are arranged on a lattice of unit length, $3\,91 \times 10^{-8}$ cm Providing therefore that molecules can be adsorbed on adjacent lattice squares it is evident that molecules so adsorbed will approach each other almost as closely as they do in a molecular collision They cannot however approach quite as closely as this because we have assumed that lateral mobility is restricted owing to the existence of potential hills between the adjacent molecules

RESULTS

The values of α so calculated are given below, together with the values of the adsorption volume calculated from x_s α is calculated from x_s by assuming that at saturation those parts of the surface which can adsorb are covered with a close-packed monomolecular layer of molecules α can alternatively be calculated by assuming one molecule is adsorbed per lattice square, and since the area of such a square is for platinum nearly equal to the square of the molecular diameter of argon, nitrogen and oxygen the difference in the two values is quite small In each case d and σ have been put equal to 4×10^{-8} cm

α FRACTION OF SURFACE AREA AVAILABLE FOR ADSORPTION

	Argon				Nitrogen			
T° K	B × 10³	α (x_s)	α (6)	α (7) and (8)	B × 10³	α (x_s)	α (6)	α (7) and (8)
77 3	5 25	1 06	0 21	0 31	—	—	—	—
83 5	—	—	—	—	5 88	0 40	0 10	0 15
90 5	5 15	0 35	0 11	0 17	5 72	0 28	0 08	0 12
112 2	4 56	0 18	0 12	0 18	5 31	0 096	0 08	0 12
170 4	4 06	0 03	0 05	0 08	4 75	0 04	0 06	0 09

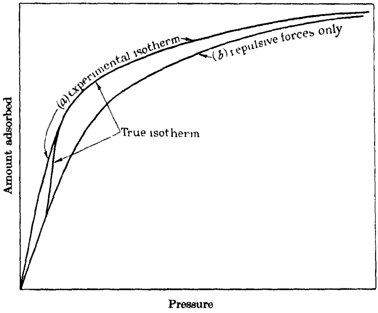


Fig 5

The agreement between the values of α calculated from b and x_s is reasonably satisfactory at 112° and 170° K having regard to the difficulty of assigning appropriate values to the virial coefficients At low temperatures the values of α calculated from b are much less than those calculated from x_s

This is to be expected if, as the shape of the adsorption isotherms suggest, intermolecular forces of attraction play an important part at these temperatures and surface concentrations in determining the molecular energy of interaction in the adsorption layer In fig 5 the true adsorption isotherm and the isotherm represented by (*a*) the experimental figures and (*b*) an adsorbed layer with only repulsive forces are drawn A brief examination of these curves shows that owing to the restricted range of the experimental data in cases where attractive forces are a significant part of the total force of molecular interaction the apparent value of b will be higher than it would be if in the adsorbed layer repulsive forces only were effective

The influence of temperature on the volume of the adsorbed phase

Hückel (1928) first pointed out that the saturation maximum as calculated from the Langmuir equation should be independent of the temperature, whereas all the experimental evidence showed that this was not the case Several authors have discussed this problem (cf Wilkins and Ward 1929, and Bradley 1931) All these have worked with x_s and have assumed that its temperature coefficient is due to some factor which alters the number of molecules which can be packed into the adsorbed phase, the

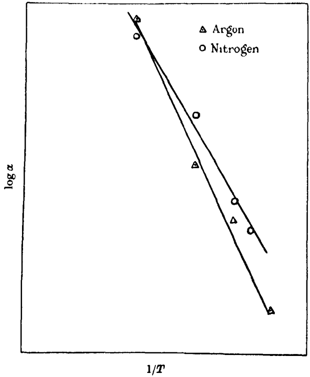


FIG 6

volume of which is independent of temperature The foregoing discussion suggests that the values of the adsorption volume (αd) calculated from x_s are the most reliable values While the values of (αd) or α so calculated cannot be of a high order of accuracy, a plot of $1/T$ against $\log \alpha$ would suggest that effects of the temperature can be represented roughly by the equation $\alpha = ce^{-\beta/T}$, where c and β are constants (see fig 6)

The magnitude of the temperature coefficient is so large that the factors discussed by Wilkins and Ward cannot possibly account for it Bradley's factor is automatically taken account of in the virial coefficient It is of course possible to argue that the sharp increase of adsorption volume with decreasing temperature is due to the formation of polymolecular layers The discussion of the previous section, while it cannot be said to dispose of this possibility entirely, does at least indicate that the results are not inconsistent with the theoretical equations which tacitly assume monomolecular layers If this view-point is correct we can draw the interesting conclusion that even on a smooth metal foil surface the surface fields are not uniform The surface can be energetically composite in two ways If we assume that the heterogeneity is associated with the type of structure postulated by exponents of the 'real Krystall' the exponential temperature effect can be explained Goetz and Hergenrother (1932) state that if N_c be the number of atoms per unit volume of a large crystal existing in perfect lattice configuration, and N_d the corresponding number of decrystallized atoms, then

$$(N_d/N_c + N_d) = f(T),$$

where $f(T)$ is exponential It is of course important to note that this equation was developed in the discussion of the behaviour of a metal near its melting point and to apply it to the problem under discussion is perhaps an unwarrantable extrapolation

There is one very important implication made in the calculation of (αd) from b It is that the molecules are concentrated together in a small fraction of the surface and not uniformly distributed over it If the adsorbed molecules are uniformly distributed over the surface area the adsorption volume, which is significant as far as the evaluation of $B_a(T)$ is concerned, is not (αd) but the total adsorption volume per sq cm, which is (d) Therefore, in order that (αd) should equal (d), for example, for argon and nitrogen at 112° K the second virial coefficient B will have to be ten times larger than the value used This is far outside the range of experimental error We are therefore driven to the conclusion that the atoms in a solid surface which are capable of adsorbing molecules are congregated together in patches and that the adsorbed molecules form islands in the surfaces The concept of islands existing in the adsorbed layer has been employed by Polanyi and Welke (1928) and Goldmann and Polanyi (1928), and also by Benton and White (1930, 1931), to explain discontinuous adsorption

The adsorption potential

In view of the difficulties which are encountered in applying the virial adsorption isotherm to the experimental results obtained at the lowest temperatures the adsorption potential has been evaluated only at 112° and 170° K , using equations (7) and (8) of the previous paper There is no evidence existing which allows us to determine σ It has been assumed that σ for the adsorbed molecule is the same as for the platinum atoms, i e $3\ 1 \times 10^{-12}$ Since the vibratory motion of the adsorbed molecules is conditioned almost entirely by the vibratory motion of the platinum atoms this is not an unreasonable assumption to make The results are tabulated below

	Nitrogen		Argon	
T° K	ϕ_7	ϕ_8	ϕ_7	ϕ_8
112 2	4400	5300	4000	4800
170 4	5700	7200	5300	6600

It appears that ϕ_7 and ϕ_8 increase markedly with temperature Actually the adsorption potential would be expected to be independent of temperature If, however, there is a change of physical state of the adsorbed layer so that melting or lateral diffusion occurs between 112 2° and 170° K , then the adsorption potential is given by ϕ_8 at 112 2° K and ϕ_7 at 170 4° K This gives us an adsorption potential independent of the temperature and

$$\phi_{N_2} \approx 5500 \text{ cals} \quad \text{and} \quad \phi_A \approx 5000 \text{ cals}$$

Lennard-Jones (1932) has developed a quantum mechanical calculation of that part of ϕ, ϕ_a which is due to the attractive forces existing between the solid surface and the adsorbed molecule He has not been able to calculate the repulsion potential ϕ_r, but he estimates it to be 40 % of ϕ_a

Now ϕ_a is given by

$$\phi_a = \frac{\epsilon^2 \bar{r}^2}{6R^3},$$

where ϵ is the electronic charge, R is the equilibrium distance of the adsorbed molecule from the surface and $\bar{r}^2$ is a quantity which can be calculated when the distribution of electrons in the adsorbed molecule is known For diamagnetic molecules $\bar{r}^2$ is given by

$$\bar{r}^2 = -\chi L\epsilon^2/6mc^2,$$

where χ is the diamagnetic susceptibility per gram atom, L is Loschmidt's number, m is the mass of the electron and c the velocity of light Following Lennard-Jones, R has been taken as the mean of the least distance of approach of the gas molecules in the solid state and of the platinum atoms in the solid state The values of ϕ are

Nitrogen	2500 cal /g mol
Argon	5300 cal /g mol

He suggests that 40 % should be allowed for the potential of the repulsive fields and this gives for ϕ

Nitrogen	1500 cal /g mol
Argon	3200 cal /g mol

The experimentally determined ϕ's are therefore much larger than these theoretical values There are two points of importance here The platinum foil was cleaned before use with hot nitric and chromic acids This treatment would probably cause the formation of a thin layer of oxide which would, of course, completely alter the surface properties It should be pointed out however, that Thomson, Stuart and Murison (1933) have shown that heating sputtered films of platinum which were covered with PtO_2 to above 225° C removed the oxide to such an extent that it could not be detected by electron diffraction Langmuir (1918, 1922) states that oxygen is so strongly and irreversibly adsorbed by platinum that it is retained even at temperatures as high as 1200° C It is possible, therefore, that the platinum was covered with a monomolecular layer of chemisorbed oxygen

A further point to be borne in mind also is that the experiments show quite clearly that although the surface is plane it is not homogeneous, and the fraction available for adsorption decreases rapidly as the temperature is raised from 80° to 190° K This is characteristic of the results of all experiments which have yet been reported where adsorption is reversible If the majority of the surface has an adsorption potential of 1500 cal /g mol for nitrogen and 3200 cal /g mol for argon the amounts adsorbed on it will be quite small compared with those parts of the surface having a potential of 5000–6000 cal /g mol , except at low temperatures where the areas of higher potential may be very nearly saturated

I should like to express my thanks to Professor F G Donnan, in whose laboratory this experimental work was carried out, for his kind interest and encouragement during the course of this investigation, and also to Mr R R M Mallock, who spent much time in the evaluation of the coefficients of the virial isotherm I am also indebted to the Department of Scientific Industrial Research for a Senior Research Award

SUMMARY

The adsorption of argon, nitrogen and oxygen on a smooth platinum foil has been measured at various temperatures between 77 and 193° K over the pressure range of 5×10^{-3} to 2×10^{-1} mm It is found that the deviations from the Langmuir adsorption isotherm are greater the lower is the temperature The values α of the fraction of the apparent surface covered by the gas at saturation is unity only for argon at 77° K In all other experiments α was much less than unity Further, the rate of decrease of α with increase of temperature is much more rapid than is given by Zeise's linear equation and can be expressed approximately by

$$\alpha = ce^{-\beta/T},$$

where c and β are constants After a discussion of the results, using the virial adsorption isotherm, the failure of the Langmuir adsorption isotherm at low temperatures is attributed to the intermolecular forces of attraction which exert a pronounced effect on the amount of gas adsorbed

The adsorption potentials of argon and nitrogen are calculated to be approximately 5000 and 5500 cal /g mol respectively These compare with the corresponding values calculated from quantum mechanics of 3200 and 1500 cal /g mol The possible reasons for the discrepancy are discussed

REFERENCES

Bawn, C E H 1932 *J Amer Chem Soc* **53**, 72–86
Benton, A F and White, T A 1930 *J Amer Chem Soc* **52**, 2325–36
— — 1931 *J Amer Chem Soc* **53**, 2807–8, 3301–14
Bradley, R S 1931 *Phil Mag* Ser 7, **11**, 690–6
Ellet, A and Zabel, R M 1931 *Phys Rev* **37**, 1102–11
Fowler, R H 1936 "Statistical Mechanics," Cambridge, 2nd edn
Goetz, A and Hergenrother, R C 1932 *Phys Rev* **40**, 643–60
Goldmann, F and Polanyi, M 1928 *Z phys Chem* **132**, 321–70
Huckel, E 1928 "Adsorption und Kapillar Kondensation," pp 150, 174

Langmuir, I 1918 *J Amer Chem Soc* **40**, 1361–1403
— 1922 *Trans Faraday Soc* **17**, 621–54
Lennard-Jones, J E 1932 *Trans Faraday Soc* **28**, 333–59
McBain, J W and Britton, G T 1930 *J Amer Chem Soc* **52**, 2198–222
Polanyi, M and Welke, K 1928 *Z phys Chem* **132** 371–83
Thomson, G P, Stuart, N and Murison, C A 1933 *Proc Phys Soc* **45**, 381–7
Wilkins, F J and Ward, A F H 1929 *Z phys Chem* **144**, 259–68
Zeise, H 1928 *Z phys Chem* **146**, 385–418

The emission band spectrum of chlorine (Cl_2^+) II

By A Elliott, Ph D, D Sc, and W H B Cameron, M Sc
Physics Department, University of Sheffield

(*Communicated by S R Milner, F R S —Received 7 October 1937*)

[Plate 8]

Introduction*

In continuation of our work (Elliott and Cameron 1937*a*) on the spectrum of chlorine excited by a high-frequency electrical discharge, we have established the fact that the discrete bands which appear are due (at least in part) to the ionized molecule Cl_2^+ In the publication referred to, we have described measurements of the wave numbers of the band heads in the region 3864–5435 A, as well as of the maxima of some continua in the ultra-violet An extensive vibrational isotope effect was identified in the discrete bands, and on the basis of this a suggestion was made that two systems were present, with origins situated at 21000 and 21340 cm $^{-1}$ respectively The vibrational analysis of Ota and Uchida (1928) was shown to be untenable, but no alternative analysis was put forward It was suggested that the intervals of 627 and 632 cm $^{-1}$, which occurred frequently, might be the vibrational frequency of the chlorine molecule in either the upper or the lower state involved

The bands considered in the present paper, with the exception of that at 26166 5 cm $^{-1}$, were all included in the first paper, but in some cases the wave numbers of the heads have been slightly amended About one-half

* Introduction rewritten and Table III revised 12 November 1937

of the ninety-four bands previously tabulated (excluding isotopes) have now been classified and assigned to Cl_2^+ The two systems previously mentioned are now identified as the subsystems of a doublet system, the system origins are much more accurately determined and are found to be situated at 20596 9 and 20797 3 cm $^{-1}$ respectively It is found that the intervals 627 and 632 (more accurately 627 9 and 633 7 cm $^{-1}$) are the vibrational frequencies of Cl_2^+ for the hypothetical levels with vibrational quantum numbers $2\frac{1}{2}$ and $1\frac{1}{2}$

PHOTOGRAPHY, MEASUREMENTS, ETC

The apparatus used for generating the chlorine and exciting its band spectrum has already been described in a previous publication (Elliott and Cameron 1937*a*) For photographing the spectrum, a Littrow spectrograph fitted with three prisms of dense flint glass was used, and arrangements for keeping constant the temperature and pressure of the surrounding air were employed These arrangements have been described elsewhere (Elliott and Cameron 1937*b*)

In the visible region the spectrum was bright enough to enable Wellington Ortho Process plates to be employed, the exposure times varying from 1 up to 6 hr In the near ultra-violet region, Ilford Double X-Press plates were used, as the exposure times would have been too long with the slower plate Even so, an exposure of 50 hr was necessary to record the band at 26166 cm $^{-1}$ In some of the last photographs taken on Double X-Press plates, a great decrease in graininess was obtained by using a fine-grain developer (Morgan, 1936)

The iron arc was used to provide a comparison spectrum, and the wave numbers of the chlorine band lines were obtained by linear interpolation between two selected iron lines These wave numbers were corrected by reference to an error curve, constructed from measurements of as many suitable iron-arc lines as could be found between the selected reference lines Usually a separate error curve was drawn for each band Great care was taken in the construction of this error curve, since a faulty curve leads to systematic errors in the wave numbers of the lines measured

DESCRIPTION OF THE BANDS

In no case is the rotation structure of a band resolved right up to the head, but many bands have been found where the structure is sufficiently resolved to show that there are two branches of approximately equal intensity, and that no others of appreciable strength are present In several bands only a single branch is apparent This is presumably due to the superposition of the two branches which occurs when the ratio of the rotation constants of upper and lower states has a suitable value (Jevons 1932 *a*)

The vibrational isotope effect has already been identified (Elliott and Cameron 1937 *a*), and it has been found possible to analyse the rotational structure of the isotopic pairs $(^{35}Cl_2)^+$ and $(^{35}Cl\,^{37}Cl)^+$ in a number of cases Plate 8 shows three such pairs

It was expected that the lines of the bands due to $(^{35}Cl_2)^+$ would show alternating intensities with a ratio of strong to weak line intensities of 1 4 to 1, since this ratio has been found in the Cl_2 absorption bands (Elliott 1930) No alternation of intensities has been found, however The evidence for this, and the explanation offered, will be given in the following section

INTENSITY MEASUREMENTS

Although none of the Cl_2^+ bands examined shows an alternation in intensity of the band lines which is visible to the eye, it was considered necessary to examine this point more closely by making microphotometer records of the densities of the band lines Bands 24308 and 23624 cm $^{-1}$ were selected for the purpose, as they are fairly free from overlying bands Reproductions of microphotometer traces of these bands are shown in Plate 8 Comparison of these traces with a calibration curve (intensity of light incident on the plate against density produced) showed that no alternation as great as 1 1 1 was present As an alternating intensity ratio of 1 4 1 had been expected, it seemed desirable to establish that the bands were really due to chlorine molecules, and not, for example, to $^{28}Si\,^{35}Cl$ and $^{28}Si\,^{37}Cl$ or to $^{31}P\,^{35}Cl$ and $^{31}P\,^{37}Cl$, molecules which would give rise to an isotopic separation similar to that expected in chlorine Conclusive evidence was given by the intensity measurements on band 25882 cm $^{-1}$ (main band and isotope), which gave a ratio of intensities for the corresponding lines in $(^{35}Cl_2)^+$ and $(^{35}Cl\,^{37}Cl)^+$ of 1 56 1 The relative abundance of these molecules calculated from the atomic weight of chlorine, 35 465, is 1 57 1 The relative abundance of $^{28}Si\,^{35}Cl$ and $^{28}Si\,^{37}Cl$, which is the same as that of $^{31}P\,^{35}Cl$ and $^{31}P\,^{37}Cl$, is 3 14 1 Whilst the close agreement of the measured intensity ratio with the relative abundance of $(^{35}Cl_2)^+$ and $(^{35}Cl\,^{37}Cl)^+$ may be in part fortuitous, there is no doubt that the carriers of the spectrum are chlorine molecules

The explanation of the absence of alternating intensities, which was suggested to us by Dr Kronig (Groningen), is probably that neither of the

TABLE I WAVE NUMBERS OF $(^{35}Cl_2)^+$ BAND LINES IN CM $^{-1}$

SUBSYSTEM I

	Band 4 → 3		Band 4 → 2		Band 6 → 2	
$J-\frac{1}{2}$	*P* branch	*R* branch	*P* branch	*R* branch	*P* branch	*R* branch
2			21666 74			
3			665 82	21668 70		
4	21036 80		664 54	668 12		
5	035 41		663 17	667 47		
6	033 83		661 62	666 74	22686 73	
7	032 27		659 99	↑	684 96	
8	030 30	21036 80	657 92		682 84	
9	028 40	↑	655 89	Super	680 70	22687 45
10	026 10		653 63	posed *P* and *R*	678 37	685 73
11	023 59		651 33	lines	675 77	683 83
12	020 96	Super	648 96		673 09	681 84
13	018 24	posed	645 93	↓	670 03	679 55
14	015 41	*P* and *R*	642 99	653 63	666 90	677 12
15	012 33	lines	640 05	651 05	663 57	674 51
16	009 05		636 60	648 45	660 03	671 57
17	005 65		633 01	645 60	656 39	668 57
18	001 98		629 34	642 61	652 42	665 30
19	20998 22		625 44	639 65	648 35	661 85
20	994 23		—	636 10	644 12	658 37
21	990 11		617 35	632 56	639 90	654 59
22	985 75		612 77	628 88	634 86	650 64
23	981 15		608 28	625 02	630 15	646 43
24	976 56		603 39	—	625 08	642 00
25	971 73		598 39	616 69	619 88	637 58
26	966 62		593 38	612 37	614 43	632 77
27	961 42		588 01	607 60	609 00	627 83
28	956 08		582 48	602 98	603 24	622 68
29	950 37		576 84	597 91		617 41
30	944 74		570 85	592 77		611 95
31	938 80		564 98	587 24		606 49
32			558 72	581 94		600 52
33				576 26		
34		↓		570 52		
35		20938 80		564 36		
36				558 12		

electronic levels involved in the production of the bands is a Σ state, and that the Λ-type doubling is too close for resolution Each line would then consist of a strong and a weak component, unresolved A somewhat similar state of affairs occurs in the $^2\Pi$–$^2\Pi$ bands of carbon, except that there the nuclear spin is zero, and the weak lines are consequently missing, which gives rise to a "staggering" of the consecutive lines in a branch

TABLE I (*continued*)

$J-\frac{1}{2}$	Band 6 → 1 P branch	Band 6 → 1 R branch	Band 8 → 1 P branch	Band 8 → 1 R branch	Band 8 → 0 P branch	Band 8 → 0 R branch
2			24306 22		24945 86	
3			305 24		944 75	24947 14
4			304 01	24307 00	943 52	946 63
5	23321 66		302 36	306 22	942 04	945 86
6	320 04		300 70	305 24	940 36	944 75
7	318 20		298 80	304 01	938 33	943 52
8	316 27	23322 24	296 74	302 63	936 21	942 04
9	314 08	320 81	294 44	300 96	933 85	940 36
10	311 55	319 07	291 99	299 14	931 39	938 52
11	309 07	317 17	289 26	297 06	928 62	936 45
12	306 26	315 03	286 30	294 86	925 67	934 18
13	303 29	312 75	283 24	292 41	922 53	931 77
14	300 10	310 24	279 93	289 73	919 22	929 06
15	296 75	307 50	276 36	286 87	915 66	926 15
16	293 13	304 68	272 66	283 85	911 78	923 00
17	289 37	301 57	268 77	280 57	907 89	919 68
18	285 39	298 30	264 67	277 14	903 65	916 14
19	281 27	294 86	260 27	273 47	899 30	912 39
20	276 96	29[illegible] 15	255 76	269 56	894 66	908 50
21	272 41	287 29	251 00	265 51	889 88	904 39
22	267 79	283 33	246 01	261 28	884 85	900 02
23	262 84	279 01	240 93	256 56	879 55	895 41
24	257 69	274 60		252 05	874 12	890 73
25	252 29	269 93		247 23	868 50	885 69
26	246 90	265 13		242 05	862 72	880 54
27	241 13	260 20				875 07
28	235 30	254 91				869 60
29	229 17	249 56				863 88
30	222 92	243 96				
31	216 48	238 10				
32	209 78	232 29				
33	203 01	226 06				
34	196 00	219 84				
35	188 86	213 27				
36	181 35	206 60				
37		199 49				
38		192 47				

CALCULATION OF THE WAVE NUMBERS OF THE BAND HEADS

Although the formulae in common use for vibrational isotope separations refer to the band origins rather than to the band heads, we have calculated the wave numbers of the heads, because these can in many cases be more accurately determined than the origins This is especially the case with the

TABLE I (*continued*)

	SUBSYSTEM I		SUBSYSTEM II		SUBSYSTEM II	
	Band 10 → 0		Band 5 → 3		Band 5 → 2	
$J-\frac{1}{2}$	*P* branch	*R* branch	*P* branch	*R* branch	*P* branch	*R* branch
2	25880 53		21346 14		21974 04	
3	879 53	25882 06	345 28		973 09	21976 38
4	878 26	881 38	344 05		971 84	975 49
5	876 70	880 53	342 74	21347 05	970 54	974 93
6	874 97	↑	341 18	346 14	969 03	974 04
7	872 98		339 48	↑	967 26	↑
8	870 84	Super	337 77	Super	965 37	Super-
9	868 47	posed	335 65	posed	963 44	posed
10	865 84	*P* and *R* lines	333 47	*P* and *R* lines	961 24	*P* and *R* lines
11	863 04		330 84		958 80	↓
12	860 05		328 32		956 12	965 37
13	856 82		325 68		953 34	963 15
14	853 35		322 83		950 39	960 77
15	849 67		319 66	↓	947 14	958 30
16	845 84		316 42	328 32	943·86	955 53
17	841 72		312 84	325 33	940 22	952 73
18	837 41		—	322 41	936 53	949 72
19	832 94		305 44	319 29	932 60	946 54
20	828 19		301 45	316 02	928 57	943 18
21	823 26		297 26	312 30	924 19	939 54
22	818 12		292 88	—	919 78	935 83
23	812 74		288 26	304 97	915 15	931 87
24			283 53	300 78	910 33	927 54
25		↓	278 64	296 70	905 38	923 38
26		812 74	273 53	292 26	900 19	918 91
27			268 21	287 64		914 16
28			262 82	282 82		909 24
29			256 96	277 83		904 11
30			251 27	272 68		899 03
31			245 02	267 27		
32			238 94	261 84		
33				256 15		
34				250 23		
35				244 07		
36				237 82		

bands of $(^{35}\mathrm{Cl}\,^{37}\mathrm{Cl})^+$, where the observations sometimes did not permit the calculation of B values, which is necessary if origin wave numbers are to be determined from the wave numbers of the band lines On the other hand, it is a fairly simple matter to determine the wave number of the head

When the D values are small (as has been found to be the case here), the mean wave numbers of adjacent P and R lines, if plotted against successive

TABLE I (*continued*)

SUBSYSTEM II

	Band 7 → 2		Band 7 → 1		Band 9 → 1	
$J-\frac{1}{2}$	*P* branch	*R* branch	*P* branch	*R* branch	*P* branch	*R* branch
3					24593 56	
4					592 34	
5	22984 28				590 89	
6	982 77	22987 55	23616 33		589 24	24593 92
7	981 04	986 46	614 65		587 37	592 67
8	979 05	985 21	612 69		585 32	591 37
9	976 92	983 63	610 41	23617 22	583 11	589 73
10	974 66	982 19	608 17	615 48	580 67	588 04
11	972 11	980 21	605 41	613 69	578 04	586 03
12	969 45	978 14	602 66	611 57	575 19	583 82
13	966 44	975 96	599 84	609 27	572 15	581 50
14	963 34	973 51	596 58	606 81	568 87	578 89
15	960 06	970 98	593 31	604 13	565 43	576 09
16	956 52	968 17	589 72	601 29	561 79	573 11
17	952 92	965 17	585 97	598 27	557 96	569 92
18	949 06	961 96	582 02	594 97	553 88	566 50
19	944 96	958 61	577 95	591 53	549 66	562 91
20	940 83	955 01	573 51	587 86	545 12	559 13
21	036 24	951 17	569 09	584 09	540 48	555 14
22	931 68	947 33	564 44	580 05	535 64	550 92
23	926 86	943 07	559 54	575 74	530 64	546 58
24	921 87	938 77	554 40	571 26	525 34	541 99
25	916 63	934 27	549 13	566 78	519 82	537 13
26		929 51	543 64	561 86	514 14	532 12
27		924 54	537 95	556 87	508 34	526 91
28		919 58	532 20	551 73	502 33	521 49
29		914 10	526 06	546 38	496 07	515 72
30			519 79	540 75	489 58	510 02
31			513 35	534 96		504 00
32			506 65	528 93		497 81
33			499 86	522 95		491 36
34				516 44		
35				509 88		
36				503 27		
37				496 92		

integers, are parabolic to a high degree of accuracy, and the vertex of the parabola is at the head of the band From such mean wave numbers of *P* and *R* lines, the vertex or head may be obtained by extrapolation This procedure was suggested to us by Mr H Bell (Manchester) The advantage of the method is that where *P* and *R* lines are not resolved (which is the case near the heads of many of the bands measured) no error is introduced because

TABLE I (*continued*)

	SUBSYSTEM II				UNCLASSIFIED	
	Band 9 → 0		Band 11 → 0		$\nu_{\text{Head}} = 21475\ 85$ cm $^{-1}$	
$J-\frac{1}{2}$	*P* branch	*R* branch	*P* branch	*R* branch	*P* branch	*R*-branch
2					21473 62	
3					472 62	
4					471 58	
5	25230 29				470 19	
6	228 64				468 82	21473 62
7	226 60		26156 90		467 15	↑
8	224 66	25230 71	154 86		465 28	
9	222 41	229 08	152 41		463 25	
10	219 94	227 23	150 00	26156 90	460 96	
11	217 27	225 29	147 18	↑	458 69	
12	214 29	223 04	144 10		456 15	
13	211 27	220 73	141 02	Super	453 51	Super
14	208 04	218 00	137 63	posed	450 66	posed
15	204 51	215 22	134 02	*P* and *R*	447 56	*P* and *R*
16	200 82	212 09	130 14	lines	444 21	lines
17	196 94	208 93	126 14		440 87	
18	192 81	205 43	121 90		437 30	
19	188 49	201 81	117 37		433 59	
20		197 97	112 73		429 51	
21		194 12	107 84		425 39	
22		189 55	102 68		421 15	
23			097 30		416 61	
24			091 82		411 95	
25			086 07		407 11	
26			080 07		402 11	
27			073 86		396 90	
28			067 45		391 50	
29			060 89		385 96	
30			054 05		380 13	
31			046 96		374 17	
32					368 07	
33				↓	361 72	
34				046 96	355 31	
35						
36						
37						↓
38						355 31

of the lack of resolution The vibrational isotope separations given in Table III have been calculated from the wave numbers of heads obtained in this way

Calculation shows that, adopting the *B* values here given, the interval between band head and origin lies between 0 55 and 0 43 cm $^{-1}$ for the bands

TABLE I (*continued*)

UNCLASSIFIED BANDS

	$\nu_{\text{Head}} = 23925\ 10$ cm $^{-1}$		$\nu_{\text{Head}} = 25243\ 14$ cm $^{-1}$		$\nu_{\text{Head}} = 25425\ 84$ cm $^{-1}$	
$J-\frac{1}{2}$	P branch	R branch	P branch	R branch	P branch	R branch
2					25423 60	
3			25239 98		422 64	
4	23920 48		238 62		421 33	
5	919 12		237 20		419 86	25423 90
6	917 47		235 56	25239 98	418 23	422 93
7	915 58	23920 93	233 52	↑	416 33	421 67
8	913 67	919 66	231 41		414 33	420 30
9	911 31	918 02	229 05	Super	412 04	418 67
10	908 91	916 24	226 60	posed P and R	409 54	416 86
11	906 32	914 31	223 80	lines	406 83	414 83
12	903 45	912 09	220 71		403 97	412 64
13	900 42	909 55	217 64		400 76	410 13
14	897 10	907 10	214 30		397 57	407 57
15	893 65	904 30	210 63		394 06	404 62
16	889 96	901 38	206 82		390 40	401 67
17	886 15	898 06	202 82		386 43	398 41
18	882 09	894 65	198 62		382 40	394 98
19	877 80	891 00	194 12		378 00	391 23
20	873 28	887 27			373 45	387 48
21	868 66	883 18		↓	368 79	383 36
22	863 76	878 91		194 12	363 89	379 02
23	858 67	874 44				374 51
24	853 29	869 87				369 99
25		864 89				364 95
26		859 81				
27		854 58				

measured The error in the vibrational isotope separations due to using head instead of origin wave numbers is therefore quite inappreciable

ANALYSIS OF THE BANDS

All the observed features of the bands are accounted for if they are assumed to belong to a $^2\Pi-^2\Pi$ system, with both states in case (a) of Hund's classification The evidence for this is as follows

The vibrational analysis shows that two systems or subsystems exist, these have slightly different vibrational constants in the upper electronic state and apparently identical vibrational constants in the lower state The rotational constants are, however, not identical for the lower electronic level, hence the systems have not a common electronic level Further, the rotational analysis (see later) shows that the rotational quantum number J is half-integral, which is the case for a doublet state (even multiplicity) The absence of alternating intensities excludes Σ states, and there remain as possibilities (if we exclude quartet states) the transitions $^2\Pi-^2\Pi$ or $^2\Delta-^2\Delta$, either of which, for Hund's case (*a*), gives two sub-bands with P and R branches of nearly equal intensity (at least for high J-values) and a very weak Q branch whose intensity falls very quickly with increasing J, and which could not be observed unless all the band lines were well resolved The states of Cl_2 being of odd multiplicity, it follows that the observed emission bands must be due to Cl_2^+, the possibility of which has been mentioned by Mulliken (1934) The lowest state of Cl_2^+ is (ibid) a $^2\Pi$ state, and it seems therefore not unlikely that the bands are $^2\Pi-^2\Pi$ (see, however, the section on the energy of dissociation of Cl_2^+) No certain evidence as to the number of missing lines is available to support this assignment In this connexion, reference should be made to the concluding remarks

The vibrational scheme proposed for $(^{35}Cl_2)^+$ is shown in Table II, the subsystems are designated I and II provisionally A striking feature is the frequent absence of alternate bands in the v' progressions Whilst all bands have been recorded which fit into these two subsystems, a transition not recorded in Table II is not necessarily absent, for the heads of some bands would fall in regions where they could not with certainty be recognized owing to the overlapping of other bands It may be noted that whereas in subsystem I the bands which have odd v' values are frequently not observed, in subsystem II it is the bands from the even levels which are occasionally missing

The intensity distribution appears to be of the type with the open Franck-Condon parabola, which is characteristic of bands for which the internuclear distances of the molecule differ considerably in the two electronic states, the upper electronic state having the greater internuclear distance Bands corresponding to the low frequency branch of the parabola would fall in the region below 18000 cm $^{-1}$ (5554 A), where the presence of strong atomic lines makes the observation of bands difficult No bands have been definitely recorded here, though they may exist

The rather meagre array of bands in Table II would not establish the correctness of the vibrational scheme, were the latter not very satisfactorily confirmed by the rotational analysis To save space, no rotational combination differences are given, but attention is directed to the table of rotation constants for $(^{35}Cl_2)^+$ (Table IV), where good agreement is found for the common level of bands in the same progression

The wave numbers of the heads of the bands of $(^{35}Cl_2)^+$ have been fitted to a formula of the usual type, in which, however, a term in v'^3 is required

$$\text{Subsystem I} \quad \nu = 20797\ 3+[572\ 3(v'+\tfrac{1}{2})-5\ 32(v'+\tfrac{1}{2})^2-0\ 013(v'+\tfrac{1}{2})^3] \\ -[645\ 3(v''+\tfrac{1}{2})-2\ 91(v''+\tfrac{1}{2})^2]$$

$$\text{Subsystem II} \quad \nu = 20596\ 9+[564\ 8(v'+\tfrac{1}{2})-4\ 13(v'+\tfrac{1}{2})^2-0\ 038(v'+\tfrac{1}{2})^3] \\ -[645\ 2(v''+\tfrac{1}{2})-2\ 89(v''+\tfrac{1}{2})^2]$$

In Table II the wave numbers of the heads calculated from these formulae may be compared with the observed values for $(^{35}\mathrm{Cl}_2)^+$ The vibrational quantum numbers have been chosen so as to give the best fit between the observed vibrational isotope separations and those calculated from equation 125 (*a*) on p 212 of Jevons' "Report on Band Spectra of Diatomic Molecules" No other numbering which we have tried gives such good agreement between observed and calculated isotope separations and the numbering adopted is therefore taken as absolute, Table III shows the observed and calculated isotopic separations, for those bands in which the heads have been calculated by extrapolation of rotation line wave numbers The agreement is in general good, though a better fit might have been expected in bands 4 → 3 and 10 → 0 of subsystem I, and in band 11 → 0 of subsystem II

Since the bands are not resolved up to the origin, it was not possible to identify *P* and *R* branches and assign rotational quantum numbers by inspection Combination differences were therefore sought which should be identical for the common level of bands in the same progression A unique set of combination differences was found, fitting the vibrational scheme of Table II The *P* and *R* branches could then be identified by noting that, of the two lines concerned in a combination difference, the *R* line has the higher frequency Since the upper state differences are given by

$$R(J)-P(J) = \Delta_2 F' = (4J+2)[B'+2D'(J^2+J+1)] = 4B'(J+\tfrac{1}{2}) \text{ approx },$$

it was possible to assign *J* values to the lines from the observed set of combination differences, as well as to calculate the rotation constant B' It was found that D' was too small to be determined The rotation constant B'' was calculated from the lower state differences using the corresponding equation Here again D'' was found to be negligible

In both states, the combination differences, when plotted against a series of integers *M*, became zero (on extrapolation) at an integral value of *M* This shows that *J* is half-integral for both states

TABLE II *a* VIBRATIONAL SCHEME FOR $(^{35}Cl_2)^+$ BAND HEADS IN CM $^{-1}$

SUBSYSTEM I

v' \ v''	0	1	2	3	4	5	6	7
0		20118 4 **20120 7**		18858 6 **18859 2**	*	*	*	*
1				†	†	*	*	*
2	21872 2 **21872 7**			19972 0 **19971 7**	19349 4 **19349 7**	†	*	*
3	†	†	†			†		*
4		22301 0 **22302·4**	21669 0 **21668 7**	21041 0 **21040 9**	20416 0 **20418 9**	†	19192 6 **19192 3**	18589 2 **18587 7**
5	†	22819 7 **22820 5**	†	21559 1 **21559 0**		20323 8 **20320 8**	†	
6	23964 4 **23967 1**	23327 4 **23327 7**	22693 9 **22694 0**		†	20828 6 **20827 9**	†	
7	†	†		†		†	20711 3 **20713 5**	†
8	24947 9 **24947 9**	24308 3 **24308 4**			†	†	†	20595 6 **20593 7**
9	†	†	24147 3 **24148 2**	†	†		†	
10	25882 7 **25883 5**	†						†

Notes to Tables II a and II b

Wave numbers of band heads have been calculated from rotation data where available

Wave numbers calculated from the band head formulae are given in heavy type

† Indicates that a band may be present (and overlaid by other bands) though there is no positive evidence for this

* Indicates that the band falls outside the range of observation

Where blank spaces occur, the bands are almost certainly absent from the photographs

TABLE II *b* VIBRATIONAL SCHEME FOR $(^{35}Cl_2)^+$

SUBSYSTEM II

v' \ v''	0	1	2	3	4	5	6
0		†			*	*	*
1	†				18589 2 **18589 6**	*	*
2	†		20383 5 **20387 5**	19758 8 **19759 6**	19137 9 **19137 5**	18522 0 **18521 1**	*
3	†	21559 1 **21560 2**	20927 2 **20926 5**	20300 1 **20298 6**	19676 1 **19676 5**	19061 2 **19060 1**	
4		†	†	20828 6 **20828 5**	†		
5	†	22609 3 **22610 8**	21976 4 **21977 1**	21348 5 **21349 2**		†	19500 4 **19500 2**
6	23762 0 **23761 4**	†	†	†			†
7		23623 9 **23623 4**	22990 2 **22989 7**	†	†	21121 6 **21123 4**	
8			†				
9	25236 3 **25235 7**	24596 8 **24596 3**	23964 4 **23962 6**			†	
10	†			†			
11	26166 5 **26167 0**	†				†	†

The rotation constants for $(^{35}Cl_2)^+$ are given in Table IV For the weaker isotopic bands of $(^{35}Cl\ ^{37}Cl)^+$ the measurements were only accurate enough to provide an approximate confirmation of the rotational isotope effect, and consequently are not reproduced here

ENERGY OF DISSOCIATION OF Cl_2^+

Using the method of Birge and Sponer, the energies of dissociation of Cl_2^+ have been determined by extrapolation of the ω-v curve, for both states of each subsystem These values, expressed in electron-volts, are given in Table V It must be remarked that all the extrapolations are very long, those for **D**′ are particularly uncertain because ω' is not linear in v'

Taking the ionization potentials of Cl and Cl_2 as 12 96 (Ruark and Urey 1930) and 13 2 V (Mackay 1924) respectively, and the energy of dissociation of the normal chlorine molecule Cl_2 as 2 468 e-volts (Jevons 1932 *b*), the energy of dissociation of normal Cl_2^+ into $^2P_{1\frac{1}{2}} + {}^3P_2$ is 2 23 e-volts This is much lower than the figure (4 4 e-volts) which we find for the lower state of the bands investigated

It is difficult to believe that the discrepancy could be entirely explained by assigning to the bands in question a lower state different from the normal one, with different dissociation products, and it is accordingly suggested that some of the numerical values given above may be in error

TABLE III. VIBRATIONAL ISOTOPE SEPARATIONS

SUBSYSTEM I			SUBSYSTEM II		
	Isotopic separation			Isotopic separation	
Band	Calc	Obs	Band	Calc	Obs
4 → 2	− 10 62 cm $^{-1}$	− 11 7* cm $^{-1}$	3 → 4	− 12 47 cm $^{-1}$	15 7* cm $^{-1}$
4 → 3	− 2 31	− 0 47	5 → 1	− 25 66	− 23 6*
5 → 1	− 25 42	− 26 7*	5 → 2	− 17 19	− 16 63
6 → 1	− 31 42	− 31 9*	5 → 3	− 8 87	− 8 43
6 → 2	− 22 94	− 22 33	5 → 6	15 13	17 9*
8 → 0	− 51 14	− 51 25	7 → 1	− 37 74	− 37 39
8 → 1	− 42 52	− 42 69	7 → 2	− 29 27	− 29 05
10 → 0	− 60 98	− 62 43	9 → 0	− 57 30	− 55 9*
			9 → 1	− 48 68	− 47 93
			11 → 0	− 66 97	− 65 04

The separations marked * are taken from the head measurements of an earlier paper (Elliott and Cameron 1937 *a*), and are on the whole less accurate than the remainder, which are derived from rotation data

TABLE IV ROTATION CONSTANTS OF $(^{35}Cl_2)^+$ (IN CM $^{-1}$)

SUBSYSTEM I			SUBSYSTEM II			UNCLASSIFIED BANDS		
Band	B'	B''	Band	B'	B''	ν_{Head}	B'	B''
*4 → 2	0 1753	0 2658	5 → 2	0 1735	0 2650	*21475 9	0 1758	0 2640
4 → 3	0 1770	0 2655	5 → 3	0 1733	0 2635	23925 1	0 1658	0 2669
6 → 1	0 1693	0 2668	7 → 1	0 1695	0 2665	*25243 1	0 1593	0 2653
6 → 2	0 1695	0 2655	7 → 2	0 1698	0 2650	25425 8	0 1660	0 2685
8 → 0	0 1645	0 2688	9 → 0	0 1668	0 2688			
8 → 1	0 1640	0 2669	9 → 1	0 1660	0 2668			
*10 → 0	0 1615	0 2693	*11 → 0	0 1598	0 2663			

* These bands have superposed *P* and *R* branches, and their rotation constants are less accurate than those of the other bands in this table

TABLE V CONSTANTS FOR ELECTRONIC ENERGY LEVELS OF $(^{35}\mathrm{Cl}_2)^+$

	ω_e	$x_e\omega_e$	$y_e\omega_e$	D*	B_0	α	r_e	ν_e
SUBSYSTEM I	cm $^{-1}$	cm $^{-1}$	cm $^{-1}$	V	cm $^{-1}$	cm $^{-1}$	A	cm $^{-1}$
Upper level	572 3	5 32	−0 013	1 67	0 186	0 0028	2 26	20797 3
Lower level	645 3	2 91	V small	4 37	0 2688	†	1 885	
SUBSYSTEM II								
Upper level	564 8	4 13	−0 038	1 62	0·183	0 0017	2 29	20596 9
Lower level	645 2	2 89	V small	4 41	0 2688	†	1 885	

* Values of the dissociation energy are derived from long extrapolations

† B_v not linear in v

Explanation of symbols

ω_e, $x_e\omega_e$, $y_e\omega_e$ coefficients of powers of $(v+\frac{1}{2})$ in the expression for the vibrational energy term

$$G_v = \omega_e(v+\tfrac{1}{2}) - x_e\omega_e(v+\tfrac{1}{2})^2 + y_e\omega_e(v+\tfrac{1}{2})^3$$

D energy of dissociation

$B_0 = \dfrac{h}{8\pi^2 cI_0}$, where I_0 is the moment of inertia of the molecule with vibrational quantum number zero

α coefficient of v in the expression $B_v = B_0 - \alpha v$

r_e internuclear distance in the equilibrium position (potential energy of the molecule a minimum)

ν_e system origin

CONCLUDING REMARKS

Besides the bands assigned to subsystems I and II, there are other bands in the spectrum which have not yet been classified, four of these bands have proved suitable for rotational analysis Choosing the numbering of the lines of these four bands so as to give rotation constants of the same magnitude as those found for the classified bands, the B values given in Table IV have been calculated These B values show that the bands cannot be assigned to either of the two subsystems, and it seems possible that they belong to another doublet system, whose origin (as shown by the magnitude of the isotopic head separations) lies close to that of the two subsystems Fragments of such a doublet system have been observed, but no analysis has yet been possible It seems likely that the molecular constants are of similar magnitude to those listed in Table V, and there is a possibility that we have here not two doublet systems, but a quartet system Some slight evidence exists that band 7 → 1 in subsystem II may be a ${}^4\Delta_{3\frac{1}{2}}$–${}^4\Delta_{3\frac{1}{2}}$ band, for the line $P(3\frac{1}{2})$ appears to be missing or of low intensity in this band (see Pl 8) However, evidence based on intensities in a crowded system such as this may be misleading The other bands are unfavourably situated for observing intensities near the origin, though in band 8 → 1 (subsystem I) the line

$P(2\frac{1}{2}) + R(5\frac{1}{2})$ appears to be weak Without higher resolution than we have at our disposal, greater certainty in this matter cannot be obtained

SUMMARY

The emission bands of chlorine have been photographed under high dispersion, and the wave numbers of many of the band lines have been measured Vibrational and rotational analyses have been made for several bands, and it is shown that these bands belong to a doublet system, possibly $^2\Pi$–$^2\Pi$, and that the carriers of the spectrum are ionized chlorine molecules No alternation of intensity is observed, showing that the Δ-type doubling is unresolved

Values for the chief molecular constants of $(^{35}Cl_2)^+$ are given

Many of the bands observed do not fit into the doublet system, and it seems likely that a second system, of similar structure to the one analysed, has its origin situated near the first one

The energy of dissociation of the lower state of the doublet bands, as determined by the method of Birge and Sponer, is considerably greater than that calculated for the normal state of Cl_2^+ from molecular and atomic ionization potentials and from the heat of dissociation of the normal state of neutral Cl_2

REFERENCES

Elliott, A 1930 *Proc Roy Soc* A, **127**, 638
Elliott, A and Cameron, W H B 1937*a* *Proc Roy Soc* A, **158**, 681
— — 1937*b* *J Sci Instrum* **14**, 28
Jevons, W 1932*a* "Report on Band Spectra of Diatomic Molecules", p 44 London The Physical Society
— 1932*b* "Report on Band Spectra of Diatomic Molecules", p 193 London The Physical Society
Mackay, C A 1924 *Phys Rev* **24**, 319
Morgan, W W 1936 *Astrophys J* **83**, 254
Mulliken, R S 1934 *Phys Rev* **46**, 549
Ota, Y and Uchida, Y 1928 *Jap J Phys* **5**, 53
Ruark, A and Urey, H C 1930 "Atoms, Molecules and Quanta", p 280 New York McGraw Hill Book Co

DESCRIPTION OF PLATE

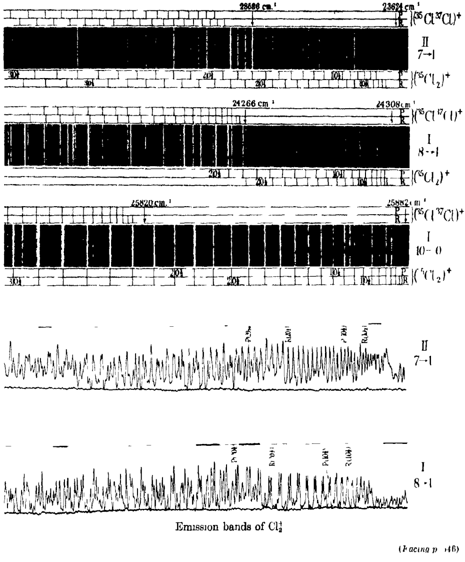
(*a*) High dispersion photograph of a portion of the emission band spectrum of chlorine (excited by a high frequency electrical discharge) showing the bands of $(^{35}Cl_2)^+$ and $(^{35}Cl\,^{37}Cl)^+$ in subsystem II, for the vibrational transition 7 → 1

(*b*) The same, for the vibrational transition 8 → 1 in subsystem I

(*c*) The same, for the vibrational transition 10 → 0 in subsystem I

(*a′*) Microphotometer record of (*a*)

(*b′*) Microphotometer record of (*b*)



Emission bands of Cl_2^+

On the solution of the laminar boundary layer equations

By L Howarth, M A, Ph D
King's College, Cambridge

(*Communicated by L Bairstow, F R S —Received* 20 *October* 1937)

Introduction

For some years after its suggestion an approximate method of solution of the boundary layer equations due to Kármán and Pohlhausen was thought to be reasonably accurate The present writer (1934) recommended it for general use because it agreed with experiment as far as the point of separation for the flow past a circular cylinder (when the observed pressure distribution was used in the theoretical solution) There seems to be little doubt that this method gives a reasonably accurate solution in a region of accelerated flow, but more recently its adequacy in a region of retarded flow has been questioned The flow past a circular cylinder is not an exhaustive test for a retarded region because the pressure rises very rapidly from its minimum value leaving little doubt as to the position of separation

Schubauer (1935) has measured the pressure distribution around an elliptic cylinder of fineness ratio 2 96 1 and also observed, by introducing smoke just beyond the separation point, the actual position of separation On applying Pohlhausen's method to his observed pressure distribution Schubauer fails to find any separation at all By measurements of the velocity distribution in the boundary layer he finds that Pohlhausen's method agrees reasonably with the observed one up to a point about five-sevenths of the way between the pressure minimum and the observed point of separation, the calculated distribution then diverges from the observed one

Fairly recently Kármán and Millikan (1934) put forward another approximate means of solution of the boundary layer equations Millikan (1936) has applied their method to Schubauer's pressure distribution, he finds separation about six-sevenths of the distance from the pressure minimum to the observed point of separation More precisely, if x is the ratio of the distance measured along the surface from the forward stagnation point to the length of the minor axis, the pressure minimum occurs when $x = 1\ 30$, the observed point of separation is $x = 1.99$ and Millikan finds $x = 1.88$ for the position of separation.

A new approximate method of solution of the boundary layer equations is suggested below. This method, when applied to Schubauer's observed pressure distribution, gives separation when $x = 1.925$ compared with the observed value 1.99. Reasons are given below (see pp. 575 and 576) for believing that the Reynolds number of Schubauer's experiments is scarcely high enough for the boundary layer equations to be valid. There is, in fact, an appreciable pressure drop across the boundary layer. It is difficult to estimate how important this pressure drop is, but the agreement between the theoretically obtained position of separation and the observed one is probably as good as could be expected in the circumstances, the agreement is somewhat better than that obtained by Kármán and Millikan's method.

This new method of solution is also compared with the exact solution given by Falkner and Skan (1930) when the velocity distribution at the edge of the boundary layer is of the form x^{-m}, fairly good agreement is obtained.

The method suggested below developed from a comparison between the results obtained by Kármán and Millikan's method, Kármán and Pohlhausen's method and an accurate solution of a particular problem.

Kármán and Millikan applied their method, when it was first introduced, to the problem of the flow along a flat plate placed edgewise to an incident stream when a retarding pressure gradient varying linearly with the distance from the leading edge is superposed. We may write the velocity distribution at the edge of the boundary layer as

$$U = b_0 - b_1 x,$$

so that the pressure gradient is $\rho b_1(b_0 - b_1 x)$, Kármán and Millikan find that separation occurs when $x^*(= b_1 x/b_0)$ is equal to 0.102, whereas Pohlhausen's method does not give separation until $x^* = 0.156$. In order to determine which, if either, of the methods is reasonably accurate it was decided to attempt to solve the boundary layer equations accurately for this velocity distribution. This problem is discussed at length below in Part I, it admits of a solution in series of the type

$$u = \frac{b_0}{2}\{f_0'(\eta) - 8x^* f_1'(\eta) + (8x^*)^2 f_2'(\eta) - \dots\},$$

where η and some of the coefficients $f_0, f_1, \dots$ are defined below in equations (2), (5)–(13). Separation occurs when

$$f_0''(0) - 8x^* f_1''(0) + (8x^*)^2 f_2''(0) - \dots = 0$$

The coefficients in this series have been determined up to and including $f_8''(0)$ They are sufficient to show that unless subsequent coefficients increase enormously—and there seems to be no reason to believe they do for, apart from f_0, f_1 and f_2 which are exceptional, they decrease steadily over the range covered—the value $x^* = 0\ 102$ for separation is much too small Furthermore, $x^* = 0\ 156$ is very much too large

Unfortunately the series converges very slowly in the neighbourhood of $8x^* = 1$ and sufficient terms have not been obtained to give the point of separation However, if we assume that the higher terms in the series continue to alternate in sign and decrease (or remain constant in absolute magnitude as the earlier ones f_3, f_8 do) it is easy to show that the point of separation lies between $x^* = 0\ 119$ and $0\ 129$ An approximate method of determining the error obtained by retaining only the terms calculated gives $x^* = 0\ 120$ as the point of separation

In Part II below a method of solution of the boundary layer equations is suggested utilizing the solution of the problem of Part I by replacing the velocity distribution at the edge of the boundary layer by a polygon of infinitesimally small sides and joining on the solution in adjacent sides by making the momentum integral

$$\theta = \int_0^\infty \left(1 - \frac{u}{U}\right) \frac{u}{U}\, dy$$

continuous at the vertex Up to the pressure minimum Pohlhausen's solution may be used Alternatively, if the number of terms computed in the solution in series starting from the forward stagnation point (Howarth 1934) is sufficient to carry the solution as far as the pressure minimum this method is to be preferred in the accelerated region If this solution does not extend quite far enough it may be extended by the solution corresponding to the one given below for the retarded region, this extension is limited, because the series does not converge rapidly enough beyond $x^* = 0\ 10$

A solution in series can also be obtained when the velocity distribution at the edge of the boundary layer is of the form

$$U = b_0 - b_1 x - b_2 x^2,$$

so that, if necessary, the method outlined above could be made more accurate by replacing the actual velocity distribution at the edge of the boundary layer by a series of parabolas

PART I

THE SOLUTION OF THE BOUNDARY LAYER EQUATIONS FOR A PARTICULAR PRESSURE DISTRIBUTION

We consider, first of all, the solution in series for a velocity distribution at the edge of the boundary layer of the form

$$U = b_0 - b_1 x,$$

where b_0 and b_1 are positive constants

We assume an expansion of the form

$$\psi = b_0^{\frac{1}{2}} x^{\frac{1}{2}} \nu^{\frac{1}{2}} \{f_0(\eta) - 8x^* f_1(\eta) + (8x^*)^2 f_2(\eta) - (8x^*)^3 f_3(\eta) + (8x^*)^4 f_4(\eta) - \quad \} \tag{1}$$

for the stream function ψ, where

$$\eta = \tfrac{1}{2} y x^{-\frac{1}{2}} \nu^{-\frac{1}{2}} b_0^{\frac{1}{2}}, \quad \text{and} \quad x^* = b_1 x / b_0 \tag{2}$$

Substituting this form in the boundary layer equation

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\nu \partial^2 u}{\partial y^2}, \tag{3}$$

where

$$-\frac{1}{\rho}\frac{\partial p}{\partial x} = -b_1(b_0 - b_1 x), \tag{4}$$

and equating coefficients of the various powers of x^*, we find that

$$f_0''' + f_0 f_0'' = 0, \tag{5}$$

$$f_1''' + f_0 f_1'' - 2f_0' f_1' + 3f_0'' f_1 = -1, \tag{6}$$

$$f_2''' + f_0 f_2'' - 4f_0' f_2' + 5f_0'' f_2 = -1/8 + (2f_1'^2 - 3f_1 f_1''), \tag{7}$$

$$f_3''' + f_0 f_3'' - 6f_0' f_3' + 7f_0'' f_3 = (6f_1' f_2' - 3f_1 f_2'' - 5f_1'' f_2), \tag{8}$$

$$f_4''' + f_0 f_4'' - 8f_0' f_4' + 9f_0'' f_4 = (4f_2'^2 - 5f_2 f_2'') + (8f_1' f_3' - 3f_1 f_3'' - 7f_1'' f_3), \tag{9}$$

$$\begin{aligned} f_5''' + f_0 f_5'' - 10f_0' f_5' + 11f_0'' f_5 = {} & (10f_1' f_4' - 3f_1 f_4'' - 9f_1'' f_4) \\ & + (10f_2' f_3' - 5f_2 f_3'' - 7f_2'' f_3), \end{aligned} \tag{10}$$

$$\begin{aligned} f_6''' + f_0 f_6'' - 12f_0' f_6' + 13f_0'' f_6 = {} & (12f_1' f_5' - 3f_1 f_5'' - 11f_1'' f_5) \\ & + (12f_2' f_4' - 5f_2 f_4'' - 9f_2'' f_4) + (6f_3'^2 - 7f_3 f_3''), \end{aligned} \tag{11}$$

$$\begin{aligned} f_7''' + f_0 f_7'' - 14f_0' f_7' + 15f_0'' f_7 = {} & (14f_1' f_6' - 3f_1 f_6'' - 13f_1'' f_6) \\ & + (14f_2' f_5' - 5f_2 f_5'' - 11f_2'' f_5) + (14f_3' f_4' - 7f_3 f_4'' - 9f_3'' f_4), \end{aligned} \tag{12}$$

$$f_8''' + f_0 f_8'' - 16 f_0' f_8' + 17 f_0'' f_8 = (16 f_1' f_7' - 3 f_1 f_7'' - 15 f_1'' f_7)$$
$$+ (16 f_2' f_6' - 5 f_2 f_6'' - 13 f_2'' f_6) + (16 f_3' f_5' - 7 f_3 f_5'' - 11 f_3'' f_5)$$
$$+ (8 f_4'^2 - 9 f_4 f_4''), \quad (13)$$

(where dashes denote differentiations with regard to η) together with the boundary conditions

$$\left.\begin{array}{l} f_r = f_r' = 0 \quad \text{when } \eta = 0 \text{ for all values of } r, f_0' = 2, \\ f_1' = \tfrac{1}{4}, \quad f_2' = f_3' = f_4' = \dots = 0 \quad \text{when } \eta = \infty \end{array}\right\} \quad (14)$$

Blasius (1908), Toepfer (1912) and Goldstein (1930, p. 15) have given the solution of (5). Indeed the problem considered is a particular case of a more general problem discussed by Goldstein (1930, p. 15), though no more of the coefficients f_r are given in that paper. The method of solution for the remaining equations, which are all linear, is exactly the same as that described by the present writer in a previous paper (1934). The following results have been obtained:

$f_0''(0) = 1.328242,$ $\quad f_1''(0) = 1.02054,$ $\quad f_2''(0) = -0.06926,$

$f_3''(0) = 0.0560,$ $\quad f_4''(0) = -0.0372,$ $\quad f_5''(0) = 0.0272,$

$f_6''(0) = -0.0212,$ $\quad f_7''(0) = 0.0174,$ $\quad f_8''(0) = -0.0147.$

The velocity distribution is given by

$$u = \frac{b_0}{2}\{f_0'(\eta) - (8x^*) f_1'(\eta) + (8x^*)^2 f_2'(\eta) - \dots\}. \quad (15)$$

The functions $f_0, f_0', f_0'', \dots, f_8, f_8', f_8''$ are tabulated in Table I.

The condition for separation $(\partial u/\partial y)_0 = 0$ leads to

$$f_0''(0) - (8x^*) f_1''(0) + (8x^*)^2 f_2''(0) - \dots = 0. \quad (16)$$

From the terms calculated it will be seen that the series converges quite slowly in the neighbourhood of $8x^* = 1$, probably at least eight more terms would be required in order to determine the value of the series in (16) to three places of decimals. We can, however, set fairly close limits to the values of x^* for separation (henceforward we shall refer to this number as x_s^*) if we assume that for $r \geqslant 9$ the values of $f_r''(0)$ do not increase in absolute

TABLE I

η	f_0	f_0'	f_0''	f_1	f_1	f_1''
0 0	0 00000	0 00000	1 32824	0 00000	0 00000	1 02054
0 1	0 00664	0 13282	1 32795	0 00494	0 09705	0 92054
0 2	0 02656	0 26553	1 32589	0 01908	0 18411	0 82054
0 3	0 05974	0 39788	1 32033	0 04142	0 26116	0 72057
0 4	0 10611	0 52942	1 30957	0 07098	0 32823	0 62073
0 5	0 16557	0 65957	1 29204	0 10674	0 38532	0 52125
0 6	0 23795	0 78756	1 26637	0 14771	0 43250	0 42258
0 7	0 32298	0 91253	1 23147	0 19291	0 46989	0 32545
0 8	0 42032	1 03352	1 18666	0 24137	0 49768	0 23094
0 9	0 52952	1 14953	1 13173	0 29214	0 51621	0 14045
1 0	0 65003	1 25954	1 06701	0 34432	0 52596	0 05565
1 1	0 78120	1 36263	0 99341	0 39706	0 52759	−0 02161
1 2	0 92230	1 45798	0 91237	0 44959	0 52195	−0 08948
1 3	1 07252	1 54492	0 82582	0 50124	0 51006	−0 14633
1 4	1 23099	1 62303	0 73603	0 55143	0 49309	−0 19093
1 5	1 39682	1 69210	0 64544	0 59973	0 47231	−0 22262
1 6	1 56911	1 75218	0 55651	0 64581	0 44900	−0 24145
1 7	1 74696	1 80354	0 47151	0 68949	0 42442	−0 24817
1 8	1 92954	1 84606	0 39234	0 73069	0 39972	−0 24418
1 9	2 11605	1 88224	0 32050	0 76946	0 37588	−0 23138
2 0	2 30576	1 91104	0 25694	0 80592	0 35367	−0 21198
2 1	2 49806	1 93392	0 20208	0 84027	0 33363	−0 18827
2 2	2 69238	1 95174	0 15589	0 87273	0 31608	−0 16240
2 3	2 88826	1 96537	0 11793	0 90357	0 30116	−0 13625
2 4	3 08534	1 97558	0 08748	0 93305	0 28879	−0 11130
2 5	3 28329	1 98309	0 06363	0 96141	0 27882	−0 08861
2 6	3 48189	1 98849	0 04537	0 98889	0 27098	−0 06879
2 7	3 68094	1 99231	0 03171	1 01566	0 26496	−0 05212
2 8	3 88031	1 99496	0 02173	1 04192	0 26045	−0 03855
2 9	4 07990	1 99675	0 01459	1 06779	0 25715	−0 02785
3 0	4 27964	1 99795	0 00961	1 09338	0 25480	−0 01965
3 1	4 47948	1 99873	0 00620	1 11878	0 25315	−0 01356
3 2	4 67938	1 99922	0 00392	1 14403	0 25203	−0 00914
3 3	4 87931	1 99954	0 00243	1 16919	0 25128	−0 00603
3 4	5 07928	1 99973	0 00148	1 19430	0 25079	−0 00389
3 5	5 27926	1 99984	0 00088	1 21936	0 25048	−0 00245
3 6	5 47925	1 99991	0 00051	1 24440	0 25029	−0 00151
3 7	5 67924	1 99995	0 00029	1 26942	0 25017	−0 00091
3 8	5 87924	1 99997	0 00017	1 29443	0 25010	−0 00054
3 9	6 07923	1 99999	0 00009	1 31944	0 25005	−0 00031
4 0	6 27923	1 99999	0 00005	1 34444	0 25003	−0 00018
4 1	6 47923	2 00000	0 00003	1 36944	0 25002	−0 00010
4 2	6 67923	2 00000	0 00001	1 39444	0 25001	−0 00005
4 3	6 87923	2 00000	0 00001	1 41944	0 25000	−0 00003
4 4	7 07923	2 00000	0 00000	1 44445	0 25000	−0 00002

TABLE I (*continued*)

η	f_2	f_2'	f_2''	f_3	f_3'	f_3''
0.0	0.0000	0.0000	−0.0693	0.0000	0.0000	0.0560
0.1	−0.0004	−0.0075	−0.0816	0.0003	0.0056	0.0560
0.2	−0.0015	−0.0163	−0.0932	0.0011	0.0112	0.0560
0.3	−0.0037	−0.0261	−0.1031	0.0025	0.0168	0.0559
0.4	−0.0068	−0.0368	−0.1108	0.0045	0.0224	0.0556
0.5	−0.0111	−0.0482	−0.1156	0.0070	0.0279	0.0551
0.6	−0.0165	−0.0599	−0.1170	0.0101	0.0334	0.0541
0.7	−0.0230	−0.0715	−0.1144	0.0137	0.0387	0.0524
0.8	−0.0307	−0.0826	−0.1077	0.0178	0.0438	0.0499
0.9	−0.0395	−0.0928	−0.0965	0.0224	0.0486	0.0463
1.0	−0.0493	−0.1018	−0.0812	0.0275	0.0530	0.0414
1.1	−0.0598	−0.1089	−0.0620	0.0330	0.0569	0.0352
1.2	−0.0710	−0.1141	−0.0397	0.0388	0.0600	0.0275
1.3	−0.0825	−0.1168	−0.0152	0.0450	0.0623	0.0185
1.4	−0.0943	−0.1171	0.0100	0.0513	0.0637	0.0082
1.5	−0.1059	−0.1148	0.0346	0.0577	0.0639	−0.0028
1.6	−0.1171	−0.1102	0.0571	0.0640	0.0631	−0.0142
1.7	−0.1278	−0.1035	0.0762	0.0702	0.0611	−0.0252
1.8	−0.1378	−0.0951	0.0908	0.0762	0.0581	−0.0352
1.9	−0.1468	−0.0855	0.1005	0.0818	0.0541	−0.0437
2.0	−0.1549	−0.0752	0.1049	0.0870	0.0494	−0.0500
2.1	−0.1619	−0.0647	0.1046	0.0917	0.0442	−0.0540
2.2	−0.1678	−0.0544	0.1000	0.0958	0.0387	−0.0555
2.3	−0.1728	−0.0448	0.0922	0.0994	0.0332	−0.0547
2.4	−0.1768	−0.0361	0.0822	0.1025	0.0278	−0.0519
2.5	−0.1800	−0.0284	0.0710	0.1050	0.0228	−0.0476
2.6	−0.1825	−0.0219	0.0595	0.1071	0.0183	−0.0422
2.7	−0.1844	−0.0165	0.0484	0.1087	0.0144	−0.0363
2.8	−0.1859	−0.0122	0.0384	0.1100	0.0111	−0.0304
2.9	−0.1869	−0.0088	0.0296	0.1109	0.0083	−0.0247
3.0	−0.1876	−0.0062	0.0222	0.1117	0.0061	−0.0195
3.1	−0.1882	−0.0043	0.0163	0.1122	0.0044	−0.0150
3.2	−0.1885	−0.0029	0.0116	0.1126	0.0031	−0.0112
3.3	−0.1888	−0.0019	0.0081	0.1128	0.0021	−0.0082
3.4	−0.1889	−0.0012	0.0055	0.1130	0.0014	−0.0059
3.5	−0.1890	−0.0008	0.0037	0.1131	0.0010	−0.0041
3.6	−0.1891	−0.0005	0.0024	0.1132	0.0006	−0.0028
3.7	−0.1891	−0.0003	0.0015	0.1132	0.0004	−0.0018
3.8	−0.1891	−0.0002	0.0009	0.1133	0.0002	−0.0012
3.9	−0.1891	−0.0001	0.0006	0.1133	0.0001	−0.0007
4.0	−0.1892	−0.0001	0.0003	0.1133	0.0001	−0.0005
4.1	−0.1892	−0.0000	0.0002	0.1133	0.0001	−0.0003
4.2	−0.1892	−0.0000	0.0001	0.1133	0.0000	−0.0002
4.3	−0.1892	−0.0000	0.0001	0.1133	0.0000	−0.0001
4.4	−0.1892	−0.0000	0.0000	0.1133	0.0000	−0.0001

TABLE I (*continued*)

η	f_4	f_4'	f_4''	f_5	f_5'	f_5''
0 0	0 0000	0 0000	−0 0372	0 0000	0 0000	0 0272
0 1	−0 0002	−0 0037	−0 0372	0 0001	0 0027	0 0272
0 2	−0 0007	−0 0074	−0 0371	0 0005	0 0054	0 0272
0 3	−0 0017	−0 0111	−0 0369	0 0012	0 0082	0 0270
0 4	−0 0030	−0 0148	−0 0365	0 0022	0 0109	0 0267
0 5	−0 0046	−0 0184	−0 0358	0 0034	0 0135	0 0262
0 6	−0 0067	−0 0220	−0 0348	0 0049	0 0162	0 0254
0 7	−0 0090	−0 0254	−0 0334	0 0067	0 0187	0 0243
0 8	−0 0117	−0 0286	−0 0314	0 0087	0 0210	0 0228
0 9	−0 0147	−0 0316	−0 0289	0 0109	0 0232	0 0210
1 0	−0 0180	−0 0344	−0 0258	0 0133	0 0251	0 0186
1 1	−0 0216	−0 0368	−0 0220	0 0160	0 0268	0 0159
1 2	−0 0254	−0 0388	−0 0175	0 0187	0 0282	0 0127
1 3	−0 0293	−0 0403	−0 0124	0 0216	0 0293	0 0090
1 4	−0 0334	−0 0412	−0 0066	0 0245	0 0300	0 0049
1 5	−0 0376	−0 0416	−0 0004	0 0276	0 0303	0 0006
1 6	−0 0417	−0 0413	0 0060	0 0307	0 0301	−0 0038
1 7	−0 0458	−0 0404	0 0125	0 0336	0 0295	−0 0083
1 8	−0 0498	−0 0388	0 0186	0 0365	0 0285	−0 0125
1 9	−0 0535	−0 0367	0 0241	0 0394	0 0271	−0 0163
2 0	−0 0571	−0 0340	0 0286	0 0420	0 0253	−0 0195
2 1	−0 0603	−0 0310	0 0319	0 0445	0 0232	−0 0221
2 2	−0 0633	−0 0277	0 0339	0 0466	0 0209	−0 0237
2 3	−0 0659	−0 0243	0 0346	0 0485	0 0185	−0 0244
2 4	−0 0681	−0 0208	0 0339	0 0503	0 0161	−0 0243
2 5	−0 0700	−0 0175	0 0321	0 0518	0 0138	−0 0233
2 6	−0 0716	−0 0144	0 0294	0 0531	0 0115	−0 0217
2 7	−0 0729	−0 0116	0 0262	0 0541	0 0094	−0 0197
2 8	−0 0740	−0 0092	0 0226	0 0549	0 0075	−0 0171
2 9	−0 0748	−0 0071	0 0190	0 0555	0 0059	−0 0147
3 0	−0 0754	−0 0054	0 0155	0 0560	0 0046	−0 0123
3 1	−0 0759	−0 0040	0 0124	0 0564	0 0034	−0 0101
3 2	−0 0762	−0 0029	0 0096	0 0567	0 0025	−0 0080
3 3	−0 0765	−0 0021	0 0073	0 0569	0 0019	−0 0061
3 4	−0 0766	−0 0014	0 0054	0 0571	0 0013	−0 0046
3 5	−0 0767	−0 0010	0 0039	0 0572	0 0009	−0 0034
3 6	−0 0768	−0 0006	0 0027	0 0573	0 0006	−0 0025
3 7	−0 0769	−0 0004	0 0019	0 0573	0 0004	−0 0018
3 8	−0 0769	−0 0003	0 0011	0 0573	0 0003	−0 0012
3 9	−0 0769	−0 0002	0 0008	0 0573	0 0002	−0 0008
4 0	−0 0770	−0 0001	0 0005	0 0573	0 0001	−0 0005
4 1	−0 0770	−0 0001	0 0003			
4 2	−0 0770	−0 0000	0 0002			
4 3	−0 0770	−0 0000	0 0001			
4 4	−0 0770	−0 0000	0 0001			

TABLE I (*continued*)

η	f_6	f_6'	f_6''	η	f_6	f_6'	f_6''
0.0	0.0000	0.0000	−0.0212	2.1	−0.0347	−0.0184	0.0165
0.1	−0.0001	−0.0021	−0.0212	2.2	−0.0364	−0.0167	0.0179
0.2	−0.0004	−0.0042	−0.0212	2.3	−0.0380	−0.0148	0.0186
0.3	−0.0010	−0.0064	−0.0211	2.4	−0.0394	−0.0129	0.0186
0.4	−0.0017	−0.0085	−0.0208	2.5	−0.0405	−0.0111	0.0181
0.5	−0.0027	−0.0105	−0.0204	2.6	−0.0416	−0.0094	0.0171
0.6	−0.0038	−0.0125	−0.0197	2.7	−0.0424	−0.0077	0.0157
0.7	−0.0052	−0.0145	−0.0188	2.8	−0.0431	−0.0062	0.0140
0.8	−0.0067	−0.0163	−0.0177	2.9	−0.0437	−0.0049	0.0124
0.9	−0.0084	−0.0180	−0.0162	3.0	−0.0440	−0.0038	0.0102
1.0	−0.0103	−0.0197	−0.0145	3.1	−0.0442	−0.0028	0.0084
1.1	−0.0123	−0.0209	−0.0123	3.2	−0.0444	−0.0021	0.0066
1.2	−0.0145	−0.0221	−0.0099	3.3	−0.0445	−0.0016	0.0051
1.3	−0.0168	−0.0229	−0.0072	3.4	−0.0446	−0.0011	0.0038
1.4	−0.0191	−0.0234	−0.0042	3.5	−0.0447	−0.0007	0.0028
1.5	−0.0215	−0.0237	−0.0009	3.6	−0.0447	−0.0005	0.0021
1.6	−0.0239	−0.0236	0.0026	3.7	−0.0447	−0.0003	0.0015
1.7	−0.0263	−0.0231	0.0061	3.8	−0.0447	−0.0002	0.0010
1.8	−0.0285	−0.0223	0.0093	3.9	−0.0447	−0.0002	0.0007
1.9	−0.0307	−0.0212	0.0122	4.0	−0.0447	−0.0001	0.0004
2.0	−0.0328	−0.0199	0.0146				

magnitude as r increases and continue to alternate in sign.† (Each term in (16) after the first is then negative.)

We can then obtain limits for x_s^* by putting first $f_r''(0) = 0$ for $r \geqslant 9$ and then by putting $|f_r''(0)| = |f_8''(0)|$ for $r \geqslant 9$. In the first case we find $x_s^* = 0.129$ and in the second $x_s^* = 0.119$. It is, therefore, fairly safe to assume that x_s^* lies within the interval 0.119 to 0.129.

We can obtain an approximate answer in the following way. We may regard the terms already obtained up to and including x^6, say, as an approximate solution.‡ It appears that each of the terms f_5' and f_6', and as far as they have been obtained, f_7' and f_8' are expressible with reasonable accuracy in the form $K_r\eta e^{-\alpha\eta^2}$, where the K_r are constants (different for different f_r).§ In fig. 1 f_6' is compared with $K_6\eta e^{-\alpha\eta^2}$ with an appropriate value −0.0221 for K_6 and $\alpha = 0.1$. It is evident by trial that α retains the same value for the different functions f_r'. If we assume that all the terms after x^6 are of the form $K_r \eta e^{-0.1\eta^3}$ the velocity u at any point may be written $(u_0 + u_1)$, where

---

† It will be seen from the equations (5)–(13) that apart from f_0, f_1 and f_2, which have exceptional right hand sides, the equations defining the f_r may be considered to be of the same form, it seems reasonable therefore to assume that they continue to alternate in sign and decrease in absolute magnitude beyond $r = 8$.

‡ The terms in x^7 and x^8 were not obtained sufficiently accurately in the outer part of the boundary layer to be used here, it is, in fact, necessary to obtain $f_r''(0)$ to an accuracy considerably greater than is required for f_r'' in the outer part of the boundary layer.

§ The error incurred by expressing f_4' in this form is not great.

$$u_0 = \frac{b_0}{2}\{f_0'(\eta) - (8x^*)f_1'(\eta) + \dots + (8x^*)^6 f_6'(\eta)\} \tag{17}$$

and

$$u_1 = \frac{b_0}{2} F(x^*)\, \eta e^{-0.1\eta^3}, \tag{18}$$

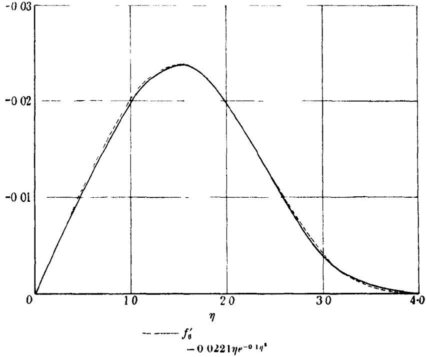


FIG 1

where $F(x^*)$ is a function to be determined. It will be noticed that the function $\eta e^{-0.1\eta^3}$ satisfies the conditions

$$u = \frac{\partial u}{\partial y} = \frac{\partial^2 u}{\partial y^2} = \dots = 0 \quad \text{when } \eta = \infty,$$

$$u = \frac{\partial^2 u}{\partial y^2} = \frac{\partial^3 u}{\partial y^3} = 0 \quad \text{when } \eta = 0,$$

a considerably greater number of conditions than the polynomial used in the ordinary Pohlhausen method. Moreover, the choice of α $(=0.1)$ so as to give agreement with the forms of f_5' and f_6' presumably corresponds to satisfying an additional condition, though it is not obvious what this condition is.

We may write the boundary layer momentum integral in the form

$$\int_0^\infty \frac{\partial}{\partial x}(u^2 - U^2)\,dy - U\int_0^\infty \frac{\partial}{\partial x}(u - U)\,dy = -\nu\left(\frac{\partial u}{\partial y}\right)_0, \tag{19}$$

i.e.

$$\int_0^\infty \frac{\partial}{\partial x}(u_0^2 - U^2)\,dy - U\int_0^\infty \frac{\partial}{\partial x}(u_0 - U)\,dy + 2\int_0^\infty \frac{\partial}{\partial x}(u_0 u_1)\,dy$$
$$+\int_0^\infty \frac{\partial}{\partial x}(u_1^2)\,dy - U\int_0^\infty \frac{\partial u_1}{\partial x}dy = -\nu\left(\frac{\partial u_0}{\partial y}\right)_0 - \nu\left(\frac{\partial u_1}{\partial y}\right)_0 \tag{20}$$

Let us write $\quad u_0 = \frac{b_0}{2}\{f_0' - 8x^* f_1' + \quad + (8x^*)^6 f_6'\} = \frac{b_0}{2}v, \qquad (21)$

$$\frac{\partial u_0}{\partial x} = \frac{b_0}{4x}\{-\eta f_0'' + 8x^*(\eta f_1'' - 2f_1') - \quad - (8x^*)^6(\eta f_6'' - 12 f_6')\} = \frac{b_0 v'}{4x},^{\dagger} \tag{22}$$

$$\frac{\partial u_0}{\partial y} = \frac{b_0^{\frac{3}{2}}}{4x^{\frac{1}{2}}\nu^{\frac{1}{2}}}\{f_0'' - 8x^* f_1'' + \quad + (8x^*)^6 f_6''\} = \frac{b_0^{\frac{3}{2}}}{4x^{\frac{1}{2}}\nu^{\frac{1}{2}}}V'^{\,\dagger} \tag{23}$$

Equation (20) may then be written

$$\int_0^\infty \{vv' + 8x^*(1 - x^*)\}\,d\eta - (1 - x^*)\left[(-\eta f_0' + f_0) + 8x^*\left(\eta f_1' - 3f_1 + \frac{\eta}{2}\right)\right.$$
$$\left. - (8x^*)^2(\eta f_2' - 5f_2) + \quad - (8x^*)^6(\eta f_6' - 13 f_6)\right]$$
$$+ F(x^*)\left[-\int_0^\infty v(\eta e^{-0\cdot1\eta^3} - 0{\cdot}3\eta^4 e^{-0\cdot1\eta^3}) + \int_0^\infty v'\eta e^{-0\cdot1\eta^3}\,d\eta - (1 - x^*)\int_0^\infty \eta e^{-0\cdot1\eta^3}\,d\eta\right]$$
$$+ 2F'(x^*)\,x^*\left[\int_0^\infty v\eta e^{-0\cdot1\eta^3}\,d\eta - (1 - x^*)\int_0^\infty \eta e^{-0\cdot1\eta^3}\,d\eta\right]$$
$$+ 2F(x^*)\,F'(x^*)\,x^*\int_0^\infty \eta^2 e^{-0\cdot2\eta^3}\,d\eta - [F(x^*)]^2\int_0^\infty \eta^2 e^{-0\cdot2\eta^3}(1 - 0{\cdot}3\eta^3)\,d\eta$$
$$= -\tfrac{1}{2}V' - \tfrac{1}{2}F(x^*) \tag{24}$$

Starting from the value of $F(x^*)$ at $x^* = 0{\cdot}0875$ given by the terms in x^{*7} and x^{*8}, equation (24) may be integrated graphically for $F(x^*)$. The condition for separation is, of course,

$$F(x^*) = -V' \tag{25}$$

---

† The dashes on v and V do not denote differentiations.

It will be seen from equation (24) that $F'(x^*)$ becomes infinite when

$$F(x^*)\int_0^\infty \eta^2 e^{-0\cdot 2\eta^2}\,d\eta+\int_0^\infty v\eta e^{-0\cdot 1\eta^2}\,d\eta-(1-x^*)\int_0^\infty \eta e^{-0\cdot 1\eta^2}\,d\eta=0. \qquad (26)$$

Once the integral curve reaches this curve it becomes imaginary. The integral curve together with the curves given by (25) and (26) is shown in fig. 2. It will be noticed that the integral reaches the curve given by (26) when $x^* = 0\cdot120$ and that it has not then reached the curve given by (25) by a quantity of the order of 2×10^{-2}.

This failure to find separation—although the actual value of the skin friction given by the last real point on the integral curve is small—must be due to the form assumed for the correction term. It seems fairly reasonable to assume that it is only in the neighbourhood of the point where $F'(x^*)$ becomes infinite that the solution is invalidated. If we suppose that the solution given is valid as far as $F(x^*) = -0\cdot110$, $x^* = 0\cdot119$ say, we may complete the solution using a result stated by Goldstein (1930, p. 4), viz.

$$\left(\frac{\partial u}{\partial y}\right)_x=\left(\frac{\partial u}{\partial y}\right)_0+\nu\frac{\left(\dfrac{\partial^4 u}{\partial y^4}\right)_0}{\left(\dfrac{\partial u}{\partial y}\right)_0}x+\nu^2\frac{\left(\dfrac{\partial^4 u}{\partial y^4}\right)^2+\dfrac{\partial u}{\partial y}\dfrac{\partial^7 u}{\partial y^7}}{2\left(\dfrac{\partial u}{\partial y}\right)^3}x^2+ \qquad (27)$$

Taking the values of $\partial u/\partial y$ and $\partial^4 u/\partial y^4$ obtained by our method of solution we find separation when $x^* = 0\cdot120$. The values of $\partial^7 u/\partial y^7$ given by the approximate method are sufficient to show that the third term in the series is negligible over the range of values of x^* $(0\cdot001)$ over which we require the solution.

An alternative method of solution is obtained by using the result given by differentiating the first boundary layer equation twice with regard to y, we find

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)=\nu\,\frac{\partial^4 u}{\partial y^4}\Big/\frac{\partial u}{\partial y}\dagger \qquad (28)$$

along the wall. We may use the skin friction given by the first nine terms of the series as an approximation and again use a term

$$u_1=\frac{b_0}{2}\phi(x^*)\,\eta e^{-0\cdot1\eta^2}$$

as a correction. Equation (28) then gives a first order differential equation for ϕ. Although we are still using the same form for the correction term, yet applying (28) (which gives the correct growth of the skin friction along the

---

† This result is, of course, contained in (27).

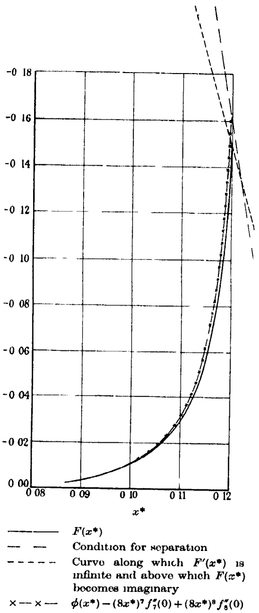

-0.18
-0.16
-0.14
-0.12
-0.10
-0.08
-0.06
-0.04
-0.02
0.00
0.08
0.09
0.10
0.11
0.12
x^*
——— $F(x^*)$
— — Condition for separation
- - - - - Curve along which $F'(x^*)$ is infinite and above which $F(x^*)$ becomes imaginary
×—×— $\phi(x^*) - (8x^*)^7 f_7''(0) + (8x^*)^8 f_8''(0)$

FIG 2

wall) instead of (24) (which is the expression of the momentum law throughout the layer) would be expected to produce a totally different result from (24) if the form assumed for the correction term were inadequate The integral curve for $\phi(x^*)$ is shown in fig 3, the corresponding value of

$$\phi(x^*) - f''_7(0)\,(8x^*)^7 + f''_8(0)\,(8x^*)^8$$

is shown dotted in fig 2 This quantity is the one we have to compare with $F(x^*)$ It will be seen that the agreement between the curves is very good and that both lead to the result that separation occurs when $x^* = 0\ 120$ The curve in fig 3 when $\phi(x^*)$ becomes infinite coincides in this case with the curve of condition for separation (Of course, the accurate value of $\partial^4 u/\partial y^4$

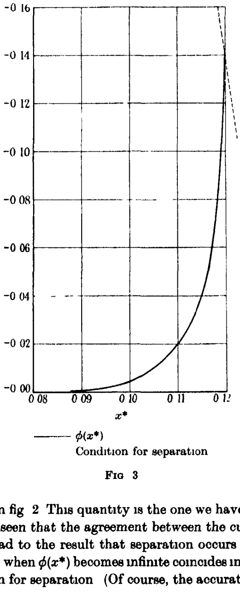


Fig 3

is zero at the separation, but the present method, being approximate, does not give this result and consequently yields an infinite value for $\phi'(x^*)$ at the point of separation)

TABLE II

x^*	$\dfrac{\nu^{\frac{1}{2}}\left(\dfrac{\partial u}{\partial y}\right)_0}{b_0 b_1^{\frac{1}{2}}}$	$\dfrac{b_1^{\frac{1}{2}}\delta_1}{\nu^{\frac{1}{2}}}$	$\chi = \dfrac{b_1^{\frac{1}{2}}\theta}{\nu^{\frac{1}{2}}}$	$\dfrac{d\chi}{dx^*}$	$\chi \Big/ \dfrac{d\chi}{dx^*}$	$\dfrac{\nu^{\frac{1}{2}}\left(\dfrac{\partial u}{\partial y}\right)_0}{U\left(-\dfrac{dU}{dx}\right)^{\frac{1}{2}}} = \dfrac{\nu^{\frac{1}{2}}\left(\dfrac{\partial u}{\partial y}\right)_0}{b_0 b_1^{\frac{1}{2}}(1-x^*)}$
0 0000	∞	0 000	0 000	∞	0 000	∞
0 0125	2 739	0 199	0 076	3 17	0 024	2 773
0 0250	1 772	0 292	0 110	2 39	0 046	1 817
0 0375	1 309	0 371	0 137	2 08	0 066	1 360
0 0500	1 011	0 447	0 162	1 93	0 084	1 064
0 0625	0 790	0 523	0 186	1 85	0 100	0 843
0 0750	0 613	0 603	0 209	1 82	0 115	0 663
0 0875	0 459	0 691	0 231	1 81	0 128	0 503
0 1000	0 315	0 794	0 254	1 84	0 138	0 345
0 1125	0 163	0 931	0 276	1 88	0 147	0 184
0 120	0 000	1 110	0 290	1 92	0 151	0 000

In Table II the values of

$$\frac{\nu^{\frac{1}{2}}\left(\dfrac{\partial u}{\partial y}\right)_0}{b_0 b_1^{\frac{1}{2}}},\quad \frac{b_1^{\frac{1}{2}}\theta}{\nu^{\frac{1}{2}}}\quad\text{and}\quad\frac{\delta_1 b_1^{\frac{1}{2}}}{\nu^{\frac{1}{2}}}$$

are tabulated against x^* over the entire range of the solution, where θ is the momentum integral $\int_0^\infty \left(1-\frac{u}{U}\right)\frac{u}{U}dy$ and δ_1 is the displacement thickness $\int_0^\infty \left(1-\frac{u}{U}\right)dy$ The values of these functions are shown graphically in figs 4, 5 and 6 The velocity distributions corresponding to $x^* = 0\ 0125$, 0 025, 0 0375, 0 05, 0 0625, 0 075, 0 0875, 0 1, 0 1125 and 0 120 are given in Table III and some of these velocity curves are shown graphically in fig 7 Values of $\nu^{\frac{1}{2}}\frac{\partial u}{\partial y}\Big/b_0 b_1^{\frac{1}{2}}$ are tabulated in Table IV for $x^* = 0\ 025$, 0 05, 0 075, 0 1 and 0 120

An additional check on the methods of solution used to complete the solution in series may be obtained in the following way If we denote by

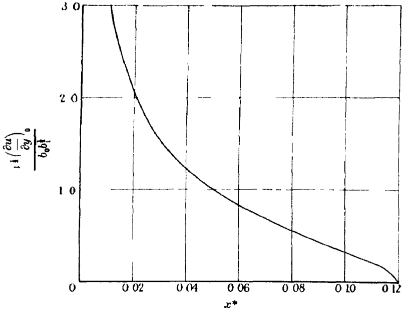

3 0
2 0
1 0
0
0 02
0 04
0 06
0 08
0 10
0 12
x^*
$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0}{b_0 b_1^{\frac{1}{2}}}$

FIG 4

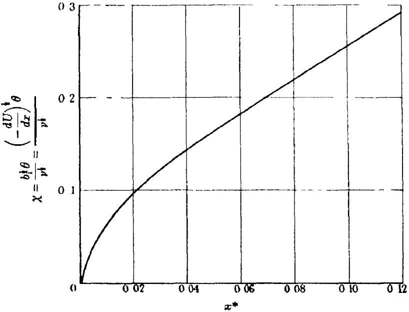

0 3
0 2
0 1
0
0 02
0 04
0 06
0 08
0 10
0 12
x^*
$\chi = \frac{b_1^{\frac{1}{2}}\theta}{\nu^{\frac{1}{2}}} = \frac{\left(-\frac{dU}{dx}\right)^{\frac{1}{2}}\theta}{\nu^{\frac{1}{2}}}$

FIG 5

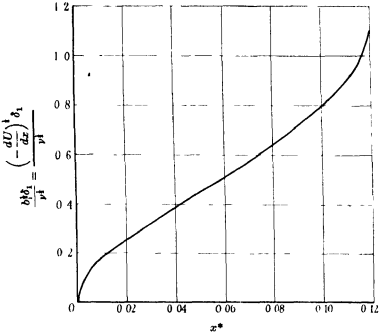

$\frac{b^{\frac{1}{2}}\delta_1}{\nu^{\frac{1}{2}}} = \frac{\left(-\frac{dU}{dx}\right)^{\frac{1}{2}}\delta_1}{\nu^{\frac{1}{2}}}$
1·2
1·0
0·8
0·6
0·4
0·2
0
0·02
0·04
0·06
0·08
0·10
0·12
x^*

Fig 6

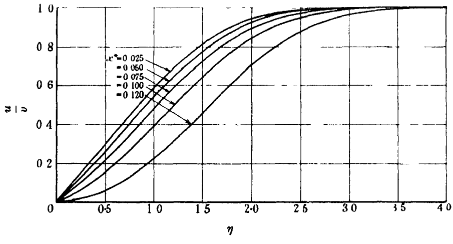

$x^* = 0·025$
$= 0·050$
$= 0·075$
$= 0·100$
$= 0·120$
$\frac{u}{v}$
1·0
0·8
0·6
0·4
0·2
0
0·5
1·0
1·5
2·0
2·5
3·0
3·5
4·0
η

Fig 7

χ the non-dimensional quantity $\theta b_1^{\frac{1}{2}}/\nu^{\frac{1}{2}}$ it is necessary in the following work to determine $d\chi/dx^*$, this may most easily be done from the momentum integral equation, which may be written

$$\frac{d\theta}{dx}+(H+2)\,\theta\frac{1}{U}\frac{dU}{dx}=\frac{\tau_0}{\rho U^2}, \tag{29}$$

where $H=\delta_1/\theta$. Equation (29) may be written, in the case of our problem, as

$$\frac{d\chi}{dx^*}-\frac{(H+2)\,\chi}{1-x^*}=\frac{\nu^{\frac{1}{2}}\left(\dfrac{\partial u}{\partial y}\right)_0}{(1-x^*)^2 b_0 b_1^{\frac{1}{2}}}. \tag{30}$$

Thus the value of $d\chi/dx^*$ corresponding to a particular value of x^* may be determined from the values shown in Table II by direct application of the momentum integral equation. If this equation were not satisfied we should expect a discrepancy to arise between the values of $d\chi/dx^*$ given by (30) and those obtained graphically from the curve or from numerical differentiation of the table of values of χ. It was in fact made very evident by trial that an error as small as 0.001 in the value of x_s^* would cause a considerable discrepancy between the value of $d\chi/dx^*$ given by (30) and that obtained from the $\chi(x^*)$ curve. Starting from the value of χ at $x^*=0.0625$ and by integrating the values of $d\chi/dx^*$ obtained from (30) we obtain the values of χ given in the second column of the following table. The values of χ already determined by integration from the velocity curves are given in the third column. The agreement will be seen to be very good.

x^*	χ Calculated from the values of $d\chi/dx^*$ given by equation (30)	χ Calculated by integration from the velocity curves
0.0625	—	0.186
0.0750	0.209	0.209
0.0875	0.232	0.231
0.100	0.254	0.254
0.1125	0.277	0.276
0.120	0.291	0.290

We may take this agreement as fairly conclusive proof that the value of x_s^* (0.120) determined by either of the approximate methods is the correct one.

PART II

NEW METHOD OF SOLVING THE BOUNDARY LAYER EQUATIONS FOR THE GENERAL CASE OF RETARDED FLOW

For simplicity we will consider in the first place the method as it was crudely conceived originally The original idea was to replace the velocity distribution at the edge of the boundary layer, in a retarded region, by a polygon of a finite number of sides We will suppose that the solution has been carried as far as the retarded region by some other method, because there seems to be little doubt that many of the existing methods (Pohlhausen's for example) are reasonably adequate in an accelerated region We may suppose, therefore, for our present purpose that the skin friction and momentum integral $\theta = \int_0^\infty \left(1-\frac{u}{U}\right)\frac{u}{U}\,dy$ are known at the commencement of the first side, we may write the velocity distribution corresponding to the first side in the form $U = b_0' - b_1 X$, where X is measured from the first vertex The essential assumption introduced now is that, once the pressure, the pressure gradient and the skin friction (or the momentum integral) are known at a particular point, the velocity distribution throughout the boundary layer is completely determined and is given by the appropriate one of the singly infinite family of velocity curves given by the problem of Part I Although not precisely true it is hoped that such an assumption offers a reasonable basis for approximation

The method consists in determining b_0 and x_0^* $(=b_1 x_0/b_0)$ so that when $x = x_0$ in the solution of the problem we have just discussed (with $U = b_0 - b_1 x$) the skin friction is identical with the known skin friction at the beginning of the first side and U is equal to b_0' there The solution at any point in the first side is then given by putting

$$x^* = x_0^* + \frac{b_1 X}{b_0},$$

the solution as far as the second vertex is therefore known and the process can be repeated The determination of b_0 and x_0^* is simple, we notice that

$$b_0' = b_0 - b_1 x_0 = b_0(1-x_0^*), \tag{31}$$

and since $(\partial u/\partial y)_{X=0}$ is supposed known we can evaluate

$$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_{\substack{y=0\\X=0}}}{b_0' b_1^{\frac{1}{2}}}$$

In virtue of (31)

$$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_{\substack{y=0\\X=0}}}{b_0' b_1^{\frac{1}{2}}} = \frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_{\substack{y=0\\X=0}}}{b_0(1-x^*)\, b_1^{\frac{1}{2}}} \tag{32}$$

We can plot the right-hand side of (32) once and for all as a function of x_0^* (see Table II and fig 8) and from the graph we can read off from the given value of

$$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_{\substack{y=0\\X=0}}}{b_0' b_1^{\frac{1}{2}}}$$

the appropriate value of x_0^* Knowing x_0^*, b_0 is determined from (31), and at any point of the first side x^* is determined by the expression $x_0^* + (b_1 X/b_0)$ Similarly x^* is determined in subsequent sides

TABLE III VALUES OF u/U

x^* η	0 0125	0 025	0 0375	0 050	0 0625	0 075	0 0875	0 100	0 1125	0 120
0 0	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000
0 2	0 125	0 117	0 108	0 099	0 089	0 078	0 066	0 052	0 034	0 010
0 4	0 251	0 237	0 222	0 205	0 188	0 168	0 146	0 120	0 085	0 038
0 6	0 377	0 358	0 338	0 317	0 293	0 267	0 237	0 202	0 152	0 085
0 8	0 498	0 477	0 455	0 430	0 403	0 372	0 337	0 294	0 234	0 149
1 0	0 611	0 590	0 567	0 541	0 513	0 480	0 442	0 394	0 325	0 227
1 2	0 711	0 692	0 670	0 645	0 617	0 585	0 546	0 498	0 426	0 318
1 4	0 796	0 779	0 760	0 738	0 712	0 682	0 646	0 598	0 527	0 416
1 6	0 864	0 850	0 834	0 815	0 794	0 769	0 736	0 692	0 625	0 517
1 8	0 914	0 904	0 891	0 877	0 860	0 839	0 812	0 776	0 716	0 616
2 0	0 949	0 942	0 934	0 923	0 910	0 894	0 872	0 844	0 794	0 708
2 2	0 972	0 967	0 962	0 954	0 946	0 934	0 918	0 897	0 858	0 787
2 4	0 985	0 983	0 979	0 975	0 969	0 961	0 951	0 936	0 908	0 853
2 6	0 993	0 991	0 990	0 987	0 984	0 979	0 972	0 962	0 943	0 903
2 8	0 997	0 996	0 995	0 994	0 992	0 989	0 985	0 978	0 967	0 940
3 0	0 998	0 998	0 998	0 997	0 996	0 995	0 992	0 989	0 982	0 965
3 2	0 999	0 999	0 999	0 999	0 999	0 998	0 997	0 994	0 991	0 981
3 4	1 000	1 000	1 000	1 000	1 000	0 999	0 999	0 998	0 995	0 990
3 6	—	—	—	—	—	1 000	1 000	0 999	0 998	0 995
3 8	—	—	—	—	—	—	—	1 000	0 999	0 998
4 0	—	—	—	—	—	—	—	—	1 000	0 999
4 2	—	—	—	—	—	—	—	—	—	1 000

Knowing x^* at every point the solution is completely determined For, in the first place, the velocity distribution is given by the corresponding one of the singly infinite system of velocity curves (some of these curves are shown in Table III and fig 7) Furthermore, if U and dU/dx are respectively the velocity and the velocity gradient at the point considered

$$U = b_0(1-x^*)$$

by the definition of b_0 and x^* Therefore

$$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0}{b_0 b_1^{\frac{1}{2}}(1-x^*)} = \frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0}{U(-dU/dx)^{\frac{1}{2}}}$$

This quantity is shown in Table II and plotted in fig 8 as a function of x^* Thus given x^*, $\dfrac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0}{U(-dU/dx)^{\frac{1}{2}}}$ is determined, and since U and dU/dx are given the skin friction can be evaluated

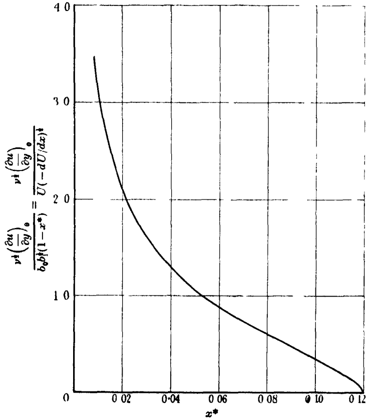


FIG 8

We naturally obtain a continuous curve for the skin friction This method is not all that could be desired, because making the skin friction continuous at the vertices corresponds to introducing an impulse at each vertex

Probably a more satisfactory method would be to make θ continuous The solution of this problem is simpler than the preceding because $b_1^{\frac{1}{2}}\theta/\nu^{\frac{1}{2}}$ is given as a function of x^* by the original problem Knowing θ initially $b_1^{\frac{1}{2}}\theta/\nu^{\frac{1}{2}}$ can be evaluated at the beginning of the first side of the polygon and the corresponding value of x_0^* read off from the graph of $b_1^{\frac{1}{2}}\theta/\nu^{\frac{1}{2}}$ against x^* The solution at any other point is given, as before, by putting

$$x^* = x_0^* + \frac{b_1 X}{b_0}, \quad \text{where} \quad b_0 = b_0'/(1-x_0^*)$$

The value of θ being thus determined at the second vertex the process may be repeated

TABLE IV TABLE OF VALUES OF $\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)\Big/ b_0 b_1^{\frac{1}{2}}$

x^*	0.025	0.050	0.075	0.100	0.120
η					
0.0	1.772	1.011	0.613	0.315	0.000
0.4	1.866	1.162	0.801	0.533	0.240
0.8	1.796	1.200	0.906	0.691	0.454
1.2	1.467	1.050	0.860	0.733	0.604
1.6	0.960	0.741	0.663	0.632	0.642
2.0	0.481	0.406	0.401	0.432	0.504
2.4	0.179	0.168	0.184	0.229	0.367
2.8	0.049	0.051	0.064	0.093	0.193
3.2	0.009	0.011	0.016	0.028	0.103
3.6	0.002	0.002	0.004	0.006	0.023
4.0	0.000	0.000	0.000	0.001	0.006
4.4	—	—	—	0.000	0.000

The objection to this method is that by making θ continuous we make the skin friction discontinuous at each join, and since the skin friction is one of the most important results of the calculation it is not satisfactory

This method of solution can, however, be extended to the case when the sides of the polygon are made to tend to zero, the number of sides tending to infinity By so doing we obtain a continuous skin friction without introducing a series of impulses

Let us approximate to the velocity distribution at the edge of the boundary layer by means of a circumscribing polygon whose sides are infinitesimally small Suppose B is one vertex of this polygon and that A and C are the points of contact of the sides through B Suppose A, B and C have abscissae x, $x+\delta x_1$, $x+\delta x$ respectively We may then write the slope of AB as $(dU/dx)_x$ and that of BC as $(dU/dx)_{x+\delta x}$

The method essentially consists in determining the value of x^* appropriate to any value of x Let us suppose the value of x^* is known at A The change δx_1^* in x^* as we pass from A to B (considered as a point of AB) is, retaining the same notation as we used previously,

$$\frac{b_1\,\delta x_1}{b_0} = \frac{b_1\,\delta x_1(1-x^*)}{b_0'} = -(1-x^*)\frac{1}{U}\frac{dU}{dx}\,\delta x_1$$

The change δx_2^* in x^* as we pass from B, considered as a point of AB, to B considered as a point of BC is given by our hypothesis that θ is continuous at B

Now $$\frac{\theta}{\nu^{\frac{1}{2}}} = \frac{\chi}{b_1^{\frac{1}{2}}}$$

Hence if θ is continuous

$$\frac{\chi(x^*+\delta x_1^*)}{\left(-\dfrac{dU}{dx}\right)_x^{\frac{1}{2}}} = \frac{\chi(x^*+\delta x_1^*+\delta x_2^*)}{\left(-\dfrac{dU}{dx}\right)_{x+\delta x}^{\frac{1}{2}}}, \tag{33}$$

i e $$1+\frac{1}{2}\left(\frac{d^2U/dx^2}{dU/dx}\right)_x\delta x = 1+\left(\frac{\dfrac{d\chi}{dx^*}}{\chi}\right)_{x^*+\delta x_1^*}\delta x_2^*\,\dagger \tag{34}$$

Moreover, $$\left(\frac{1}{\chi}\frac{d\chi}{dx^*}\right)_{x^*+\delta x_1^*}\delta x_2^* = \left(\frac{1}{\chi}\frac{d\chi}{dx^*}\right)_{x^*}\delta x_2^*,$$

retaining only first order quantities Thus

$$\delta x_2^* = \frac{1}{2}\left\{\frac{d^2U}{dx^2}\,\chi\bigg/\frac{dU}{dx}\frac{d\chi}{dx^*}\right\}\delta x \tag{35}$$

Further, the change δx_3^* in x^* as we pass from B to C along BC is given by

$$\delta x_3^* = -(1-x^*-\delta x^*)\left(\frac{1}{U}\frac{dU}{dx}\right)_{x+\delta x}(\delta x-\delta x_1),$$

---

† Since χ is known to be a differentiable function of x^*, and since δx_2^* will be small so long as the change in slope from AB to BC is small, we can write

$$\chi(x^*+\delta x_1^*+\delta x_2^*) = \chi(x^*+\delta x_1^*)+\left(\frac{\partial\chi}{\partial x^*}\right)_{x^*+\delta x_1^*}\delta x_2^*$$

to the first order Equation (34) then follows immediately where δx^* is the total change in x^* from A to C Thus to the first order

$$\delta x_3^* = -(1-x^*)\left(\frac{1}{U}\frac{dU}{dx}\right)(\delta x - \delta x_1) \tag{36}$$

Hence the total change in x^* as we pass from A to C is

$$\begin{aligned}\delta x^* &= \delta x_1^* + \delta x_2^* + \delta x_3^* \\ &= \left\{-(1-x^*)\frac{1}{U}\frac{dU}{dx} + \frac{1}{2}\frac{\chi}{\frac{d\chi}{dx^*}}\frac{d^2U/dx^2}{dU/dx}\right\}\delta x\end{aligned}$$

Proceeding to the limit when the sides all tend to zero we see that

$$\frac{dx^*}{dx} = -\frac{1}{U}\frac{dU}{dx}(1-x^*) + \frac{1}{2}\frac{\chi(x^*)}{\frac{d\chi(x^*)}{dx^*}}\frac{d^2U/dx^2}{dU/dx} \tag{37}$$

(37) is a differential equation for x^* in terms of x, since $\chi(x^*)$ is a known function of x^* shown in Table II and fig 5, and U is a known function of x The solution may be obtained graphically or otherwise once the initial value of x^* is known If, as before, we suppose the solution starts from a given value θ_0 of θ at $x = x_0$, then we can evaluate

$$\left(-\frac{dU}{dx}\right)_{x_0}^{\frac{1}{2}}\frac{\theta_0}{\nu^{\frac{1}{2}}}$$

and determine the initial value of x^* from the graph of $\chi(x^*)$ Thus the initial value of x^* corresponding to x_0 is determined, and hence the complete relation between x^* and x is known from the solution of (37) Once the value of x^* appropriate to a given value of x is known we can determine the skin friction and the velocity distribution The skin friction is obtained by reading off the corresponding value of

$$\frac{\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0}{b_0(1-x^*)\,b_1^{\frac{1}{2}}}$$

from Table II or fig 8, since $b_0(1-x^*) = U$ and $b_1 = -dU/dx$, $\nu^{\frac{1}{2}}(\partial u/\partial y)_0$ is determined The velocity distribution is given by the appropriate one of the singly infinite family of velocity curves given by different values of x^* in the exact solution of the original problem Some of these curves are shown in Table III and fig 7

So far it has been assumed that another method has been used to carry the solution just into the retarded region (i e just past the pressure minimum)

In all applications it would probably be most convenient to start the solution at the pressure minimum. Since dU/dx vanishes at the pressure minimum it is evident that x^* also vanishes there (since it contains dU/dx as a factor). Further, equation (37) has a singular point when $x^* = 0$ and $dU/dx = 0$, since $\chi \sim x^{*\frac{1}{2}}$ (see below) for small values of x^*. The appropriate value of dx^*/dx must be determined from the value of θ at the pressure minimum obtained, of course, from the method of solution used as far as the pressure minimum.

We may proceed here as we did before. Let us suppose that A is the velocity maximum (the pressure minimum) and suppose that AB and BC are two adjacent sides of the circumscribing polygon (AB will therefore be parallel to the x-axis). As before we take the abscissae of A, B and C to be x, $x+\delta x_1$ and $x+\delta x$. Thus the slope of BC is, to the first order, $(d^2U/dx^2)_x\,\delta x$. The value θ_0 of θ is supposed given at A by the method of solution used in the accelerated region and will be used as our starting point. Apart from a first order quantity the value of θ at B considered as a point of AB will be θ_0. Therefore the value of χ at B considered as a point of BC is determined by making θ continuous at B and is $\dfrac{\theta_0}{\nu^{\frac{1}{2}}}\left(-\dfrac{d^2U}{dx^2}\delta x\right)^{\frac{1}{2}}$ neglecting quantities of the order of $(\delta x)^{\frac{1}{2}}$. Now χ vanishes at A, so we may write, at C

$$\delta\chi = \frac{\theta_0}{\nu^{\frac{1}{2}}}\left(-\frac{d^2U}{dx^2}\delta x\right)^{\frac{1}{2}} \dagger \tag{38}$$

Moreover, for small values of x^*, χ is of the form $Kx^{*\frac{1}{2}}$,‡ where $K = 0.664$. Therefore to the same order

$$\delta\chi = K(\delta x^*)^{\frac{1}{2}}, \tag{39}$$

since $\chi = \delta\chi$ when $x^* = \delta x^*$. Equating the values of $\delta\chi$ from (38) and (39) we see that

$$\frac{dx^*}{dx} = -2.269\,\frac{d^2U}{dx^2}\,\frac{\theta_0^2}{\nu} \tag{40}$$

Thus, starting from the pressure minimum with a known value of θ_0 it is necessary to solve the equation

$$\frac{dx^*}{dx} = -\frac{1}{U}\frac{dU}{dx}(1-x^*) + \tfrac{1}{2}\chi\,\frac{d^2U}{dx^2}\bigg/\frac{dU}{dx}\,\frac{d\chi}{dx^*}$$

---

† Since the change in χ from B to C is $0(\delta x)^2$

‡ From the solution in series for u we see that

$$\chi = \frac{b^{\frac{1}{2}}\theta}{\nu^{\frac{1}{2}}} = 2x^{*\frac{1}{2}}\int_0^\infty \tfrac{1}{2}f_0(1-\tfrac{1}{2}f_0)\,d\eta + 0(x^{*\frac{3}{2}})$$

with the boundary conditions $x^* = 0, \dfrac{dx^*}{dx} = -2\ 269 \dfrac{d^2U}{dx^2}\left(\dfrac{\theta_0^2}{\nu}\right)$ (the additional boundary condition being required since $x^* = 0$ is a singular point with an infinite number of integrals passing through it) The function $\chi \Big/ \dfrac{d\chi}{dx^*}$ is shown in Table II and fig 9, tabulated or plotted against x^*

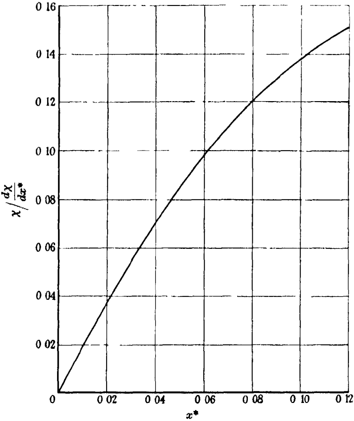


FIG 9

COMPARISON WITH A KNOWN SOLUTION

Falkner and Skan have given the exact solution of the boundary layer equation for a velocity distribution $U = x^{-m}$ at the edge of the boundary layer for values of m between 0 09 and 0 (they also give the solution for negative values of m, but that need not concern us here) The present method of solution does not lend itself to starting from the infinite value of U at

$x = 0$, we can, however, start from the exact value of θ at any point further downstream and continue by means of the present solution in order to test the adequacy of this method of solution Equation (40) reduces to

$$\frac{dx^*}{dx} = \frac{1}{x}\left[m(1-x^*) - \tfrac{1}{2}(m+1)\frac{\chi}{\dfrac{d\chi}{dx^*}}\right] \tag{41}$$

Falkner and Skan do not, however, give sufficient information to enable us to calculate θ at any point for a given value of m, but we may choose the initial value of x^* for starting our solution so that it leads to the right skin friction

For the value $m = 0.09$ Falkner and Skan find a solution in which the skin friction vanishes for all values of x Let us consider the same problem and start our solution from the value of x^* which gives the right skin friction initially (zero), i e we start from $x^* = 0.120$

With $x^* = 0.120$ we find from (41) that

$$\frac{dx^*}{dx} = \frac{1}{x}\left[m(0.880) - \frac{m+1}{2}\,0.151\right]$$

In order that the skin friction should vanish everywhere dx^*/dx must vanish everywhere so that x^* retains the value 0.120 for all values of x, we require therefore

$$\frac{m}{m+1} = \frac{0.0755}{0.880},$$

which leads to a value $m = 0.0938$

Thus the present method gives a value 0.0938 for m compared with the value 0.09 obtained by Falkner and Skan Hartree, using the differential analyser, has recently given a value 0.0905 for this quantity (1937)

Pohlhausen's equation for $\lambda\left(=\dfrac{dU}{dx}\dfrac{\delta^2}{\nu}\right)$ may be put in the form

$$\frac{d\lambda}{dx} = \frac{1}{U}\frac{dU}{dx}f(\lambda) + \frac{1}{\dfrac{dU}{dx}}\frac{d^2U}{dx^2}h(\lambda) \quad \text{(Howarth 1934, p 14)}$$

This method gives separation when $\lambda = -12$, so that applying a similar analysis here to determine the appropriate value of m we find

$$\frac{m}{m+1} = -\frac{h(-12)}{f(-12)},$$

which leads to a result $m = 0.100$

The agreement given by the present method with the exact solution is much better than that given by Pohlhausen's method and may, I think, be considered quite good for the following reason. The problem to which it has been applied is probably the one most likely to show up any defects because the pressure gradient is decreasing in the direction of the flow. It is evident from the values already given that with this type of flow the present method gives separation too late, an examination of equation (37) suggests that when the pressure gradient is increasing in the direction of the flow (as it usually does in practical cases between pressure minimum and separation) the present method may give separation, if anything, too early since the sign of the last term in the equation depends on the curvature of the pressure curve.

Returning now to the method suggested in this paper we can test it against the exact solution by choosing some small value of m, say, 0.04. Then

$$\frac{dx^*}{dx} = \frac{1}{x}\left[0.04(1-x^*) - 0.52\frac{\chi}{\dfrac{d\chi}{dx^*}}\right] \tag{42}$$

With this value for m the graph given by Falkner and Skan shows that

$$\nu^{\frac{1}{2}}\left(\frac{\partial u}{\partial y}\right)_0 = 0.26x^{-\frac{(3m+1)}{2}}\ \dagger \tag{43}$$

Starting from the value $x = 1$ we see that

$$\nu^{\frac{1}{2}}\frac{\partial u}{\partial y}\Big/ U\left(-\frac{dU}{dx}\right)^{\frac{1}{2}} = \frac{0.26}{0.2} = 1.3$$

The appropriate value of x^* determined from fig. 8 is 0.040. When $x = 1$ and $x^* = 0.040$, $dx^*/dx = 0.015$. Now $dx^*/dx = 0$ for all values of x when $x^* = 0.042$. It is made clear by trial that the solution is given by $x^* = 0.042$ for all values of x apart from a small initial region of the order $x = 0.2$ when x^* passes from 0.040 to 0.042. Now when $x^* = 0.042$,

$$\nu^{\frac{1}{2}}\frac{\partial u}{\partial y}\Big/ U\left(-\frac{dU}{dx}\right)^{\frac{1}{2}} = 1.25,$$

so that

$$\nu^{\frac{1}{2}}\frac{\partial u}{\partial y} = 0.25x^{-\frac{(3m+1)}{2}}, \tag{44}$$

for all values of x.

---

† This value was obtained (more precisely it was half this value) from a small scale graph, so that its accuracy is somewhat doubtful.

Comparing (43) and (44) we see that the agreement may be considered excellent especially in view of the determination of the 0 26 in equation (43) from a small scale graph It is, in fact, almost impossible to distinguish between the abscissae corresponding to the values 0 25 and 0 26

COMPARISON WITH EXPERIMENT

Schubauer (1935) has measured the pressure distribution around an elliptic cylinder of fineness ratio 2 96 1 It is doubtful, however, whether the Reynolds number he used was sufficiently high

If V is the velocity at infinity, if u is the ratio of the velocity at a point in the boundary layer to V and if p is the ratio of the pressure to $\frac{1}{2}\rho V^2$ the large terms in the second boundary layer equation (the equation of flow perpendicular to the wall) are

$$\frac{\partial p}{\partial y} = -\frac{2u^2}{r},$$

where y is the ratio of the normal distance from the boundary to the minor axis B and r is the ratio of the radius of curvature of the boundary at the point considered to B Therefore the difference between the value of p at a point y and the value p_s of p at the surface is given by

$$p - p_s = -\frac{2U^2\delta}{r}\int_0^{y/\delta}\frac{u^2}{U^2}d\left(\frac{y}{\delta}\right),$$

where U is the value of u and δ is the value of y at the edge of the boundary layer

Now in Schubauer's measurements, when $x = 1\ 946$ (he found separation at 1 99), the value of y given by $u = 0\ 95U$ is approximately 0 04 so that the change in p across the boundary layer is approximately $\frac{2 \times 1\ 5 \times 0\ 04}{4} \times \frac{1}{3}$, i e 0 010 The value 1/3 which has been used here for the integral is slightly low compared with the one obtained by using Schubauer's velocity distribution (a value 0 4 was obtained by the present writer by numerical integration), Schubauer also gives the result that r varies between 0 17 at the end of the major axis to 4 4 at the end of the minor axis ($x = 1\ 6$), so that the value 4 for r at a point well past the end of the minor axis is probably on the large side Thus the value 0 010 is a very conservative estimate of the change in p across the layer Now the third figure in p is significant in determining the pressure gradient to the accuracy with which it is required in the retarded region The average pressure gradient over the region from the pressure minimum ($x = 1\ 3$) to the observed point of separation ($x = 1\ 99$) is 0 190 approximately Moreover the change in p across the layer is not a constant but depends primarily on δ so that the pressure distribution at the edge of the boundary layer probably gives an entirely different pressure gradient distribution (Schubauer estimates that δ does not vary greatly in the retarded region, he gives a variation of from 0 040 to 0 049, but Millikan who has also examined Schubauer's experimental data gives a considerably greater variation from 0 032 and 0 051) There is therefore a reasonable doubt whether the pressure gradient obtained by Schubauer is adequate for the purpose of testing the accuracy of the method of solution suggested above

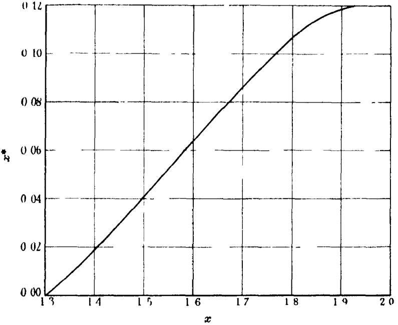


FIG 10

However, using the velocity distribution at the edge of the boundary layer deduced by Schubauer from his observed pressure distribution, I estimated dU/dx and d^2U/dx^2 graphically in the retarded region and applied the method suggested above starting from the value of θ at the pressure minimum given by Pohlhausen's method The integral curve for x^* is shown in fig 10, it will be noticed that the value of x at the point of separation is 1 925 and that x^* does not rise much above the separation value (since dx^*/dx is small at the point of separation) The values of the skin friction in the retarded region are shown graphically in fig. 11.† Schubauer using smoke to locate the separation point found $x = 1.99$ for separation, whilst Millikan using Kármán and Millikan's method gives $x = 1.88$.

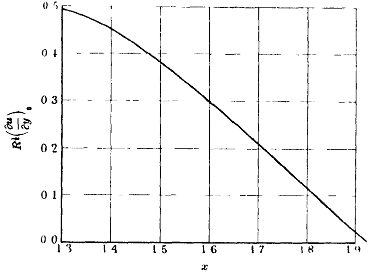


FIG. 11

It should be remarked, however, that using the values Schubauer estimated for dU/dx and d^2U/dx^2 the present method failed to give separation, x^* did not rise above 0.10. The values estimated by Schubauer and those I used are shown in Tables V*a* and V*b* respectively. The reason for failure was largely the very early inflexion in the velocity curve given by Schubauer.

TABLE V*a*

x	U	$\frac{dU}{dx}$	$\frac{d^2U}{dx^2}$
1.250	1.295	0.002	−0.170
1.350	1.295	−0.012	−0.183
1.457	1.292	−0.033	−0.192
1.600	1.284	−0.070	−0.208
1.700	1.275	−0.094	−0.203
1.832	1.261	−0.121	−0.012
1.900	1.252	−0.120	+0.108

---

† It is impossible to estimate at all accurately the skin friction from Schubauer's measurements. In addition to the large correction he finds it necessary to apply near to the surface, the points are scattered in the retarded region. I have not attempted, therefore, to include his results.

TABLE V*b*

x	U	$\frac{dU}{dx}$	$\frac{d^2U}{dx^2}$
1.3	1.295	0.000	−0.215
1.4	1.294	−0.024	−0.255
1.5	1.290	−0.050	−0.260
1.6	1.284	−0.075	−0.255
1.7	1.275	−0.100	−0.220
1.8	1.264	−0.119	−0.140
1.9	1.252	−0.126	−0.005
1.925	1.249	−0.126	+0.055

and for which I have been unable to find any evidence. My estimation gave a value 1.90 for x at the inflexion whilst Schubauer's value was about 1.84. In any case the comparatively large value of the pressure difference across the layer makes it very doubtful whether we are justified in determining the second derivative at all.

It may also be remarked that when Pohlhausen's method was applied to the values I estimated for dU/dx and d^2U/dx^2 the maximum value for $-\lambda$ was about 7.5, a value considerably below the value 12 required by this method for separation.

I wish to express my gratitude to the Air Ministry for providing me with a computer to perform much of the mechanical labour necessary in the solution of the equations (7)–(11).

SUMMARY

Part I. The problem of the flow along a flat plate placed edgewise to a steady stream, when a retarding pressure gradient varying linearly as the distance x from the leading edge of the plate is superposed is discussed. If y denotes distance measured perpendicular to the plate, a solution is obtained in the form of a power series in x whose coefficients are functions of $y/x^{\frac{1}{2}}$. Differential equations are obtained for these coefficients. Seven of the coefficients have been obtained with reasonable accuracy and the eighth and ninth roughly. Unfortunately it appears that about eight more terms are required to carry the solution to the point of separation, the work involved in their determination is prohibitive. Two approximate methods have been developed for determining the error when the first seven terms of the series are used as an approximation. These methods lead to the determination of the point of separation and are in agreement as to

its position If b_0 is the velocity at the edge of the boundary layer at the leading edge of the plate and b_1 is the velocity gradient, separation is found when $b_1x/b_0 = 0\ 120$

Part II A method is developed for the solution of the boundary layer equations in any retarded region It is obtained by replacing the velocity distribution at the edge of the boundary layer by a circumscribing polygon of infinitesimal sides and applying the preceding solution to each of these sides, making the momentum integral continuous at each vertex The problem is thereby reduced to the solution of a first order differential equation

REFERENCES

Blasius 1908 *Z Math Phys* **56**, 1–37
Falkner and Skan 1930 *Reports and Memoranda*, No 1314
Goldstein 1930 *Proc Camb Phil Soc* **26**, 1
Hartree 1937 *Proc Camb Phil Soc* **33**, 237
Howarth 1934 *Reports and Memoranda*, No 1632
Kármán and Millikan 1934 *Nat Advis Com Aero*, Report No 504
Millikan 1936 *J Aero Sci* **3** (January), 91–4
Schubauer 1935 *Nat Advis Com Aero*, Report No 527
Toepfer 1912 *Z Math Phys* **60**, 397–8

The crystal structure of insulin

I The investigation of air-dried insulin crystals

By Dorothy Crowfoot

Department of Mineralogy and Crystallography, Oxford

(*Communicated by R Robinson, F R S —Received* 11 *November* 1937)

[Plate 9]

The crystal structure of insulin is of interest not only in itself but as part of the wider problem of protein structure considered as a whole The early X-ray work on protein fibres gave strong support to the original chemical theories of the existence of polypeptide chains as the essential backbone of the protein molecule The comparatively small spacings of 3–4, 4 5 and 10–11 A which are most prominent on a large section of X-ray photographs of the scleroproteins are easily interpretable in terms of a fundamental chain structure (Astbury and Woods 1933, Astbury and Sisson 1935) It is on the other hand difficult to reconcile a chain structure with certain of the physical properties of the globular proteins, particularly their crystalline form and their behaviour in the ultracentrifuge as approximately spherical masses of molecular weight some multiple of 35,000 (Svedberg 1937 *a*) Yet there is much evidence, physical as well as chemical, that the two types of protein structure are closely interrelated Certain of the natural protein fibres show X-ray reflexions corresponding to very long spacings superimposed on the simple β keratin type of fibre pattern These long spacings—up to 300 A in feather keratin (Astbury and Marwick 1932), tendon and collagen (Wyckoff and Corey 1936)—are more of the order of magnitude of the spacings which define the crystal structure, i e intermolecular distances, in crystals of the soluble proteins And further fibre forms can be obtained from the crystalline proteins, particularly relevant observations being those on edestin and excelsin (Astbury, Dickinson and Bailey 1935) Several attempts have been made on the theoretical side to reconcile these facts, and two views of protein structure in particular have been stated in some detail Astbury has advocated a variety of straight chain structures for the different fibre proteins which may be coiled into spiral forms or folded in layers in the globular proteins (Astbury and others 1935), as also discussed by Pauling and Mirsky (1936), while D M Wrinch (1937 *a*) has shown that the occurrence of an internal cyclizing process might lead to the development from polypeptide chains of beautifully symmetrical "cyclol" networks both open and space enclosing But on the experimental side many more observations are necessary before even these quite different hypotheses can be distinguished, and one possible approach seems to be through a more complete X-ray study of the crystalline globular proteins

In recent years it has been possible to obtain X-ray diffraction patterns from a considerable number of protein crystals In most cases these patterns have been powder photographs only and can give little more than a record that some degree of orderly arrangement of molecules exists in the crystals (Fankuchen 1934, Wyckoff and Corey 1935, Magnus-Levy, Meyer and Lotmer 1936) So far in three examples only, pepsin (Bernal and Crowfoot 1934 *a*), insulin (Crowfoot 1935) and excelsin (Astbury and others 1935), have actual single crystal X-ray diffraction patterns been obtained, and in these it is possible to carry the deductions from the X-ray analysis a stage further The first photographs permitted the calculation of the actual crystal unit cell size in each case and hence the order of magnitude of the protein molecular weight The crystallographic method at this stage cannot fix the exact submultiple of the cell molecular weight which is the true protein molecular weight unless controlled by chemical or other evidence In these particular three protein crystals the weights found agree very well in order of magnitude with those measured by Svedberg, provided that in pepsin there are six Svedberg units in the crystal unit cell, in insulin and excelsin only one It seems at present simplest to refer to the weights found by Svedberg's method as (approximately) the weights of the protein molecules themselves, i e to the Svedberg units as the protein molecules, leaving further work to establish possible subdivision

The coincidence, that in the case of insulin and excelsin the unit cell molecular weight is equal to the protein molecular weight given by the ultracentrifuge method, has direct implications for the internal structure of the protein molecules themselves In the first place the symmetry of the molecule must be that of the crystal at least to the degree of accuracy measured by the available X-ray spectra This point alone is worth further investigation, since the preliminary data indicated trigonal symmetry for both insulin and excelsin Secondly, any distribution of scattering matter within the unit cell which can be deduced from the X-ray intensities will be that within the actual Svedberg molecule It is naturally impossible in the case of exceedingly complex structures such as these to attempt to apply directly ordinary trial and error methods to the analysis of the intensities of the X-ray reflexions, but the recent introduction of Patterson analysis (Patterson 1934, 1935) does supply at least the possibility of a description of the intensity variations in terms of prominent interatomic distances within the structure This method could be applied to study any protein crystal from which sufficiently good X-ray reflexions are obtainable, and of the three so far examined insulin seemed the most hopeful for a trial attempt, since the crystal structure is simpler than that of pepsin and the molecular structure simpler, at least in order of magnitude, than that of excelsin The preliminary data published in 1935 on X-ray diffraction from single crystals of insulin showed both that the variation of the X-ray intensities was perfectly definite and characteristic and also that there was, to the eye, no apparent relation between these and any protein fibre pattern

The application of the methods of X-ray crystallography to the problem of the crystal structure of proteins has been much hindered experimentally by two factors First, most of the beautiful protein crystals described in the literature contain water of crystallization which they lose very readily on exposure to the air Secondly as work on simpler organic molecules has shown, the more complicated is the molecular structure, the smaller tends to be the proportion of atoms contributing to any one reflecting plane (Robertson 1934) The absolute reflecting power of protein crystals for X-rays is therefore generally very weak As far as the first factor is concerned, it has proved possible in the case of insulin to obtain diffraction effects from ordinary air-dried crystals The percentage of water of crystallization in such crystals is small and this is not lost on exposure to the air On the contrary, the vacuum-dried crystals prove very hygroscopic and rapidly regain to a considerable degree their former weight The second factor is probably that mainly responsible for the failure of the earlier attempts to obtain crystalline diffraction effects from insulin, e g those of W H George (1928), of K Freudenberg and his co-workers (1930) and also of H Sims and D A Scott (1930) These authors all examined crystalline powders, from which they could obtain no more than an amorphous ring at about 4 5 A, together with diffuse scattering extending out from the centre of the photograph to a limit of about 10 A

This diffuse scattering covers in extent the area of the true crystal diffraction effects obtained in 1935 It is very difficult to distinguish actual powder lines in this area even on photographs exposed for from 30–40 hr and quite impossible to base any crystal structure determination on the lines that can be perceived For this reason alone one might hesitate to accept Clarke and Corrigan's report on the crystal structure of insulin based on the finding of long spacings in 1932 The structure proposed does not agree with that given in this paper and no later confirmation of the measurements has appeared * In order to observe discrete and interpretable diffraction effects it does seem most desirable, if not always absolutely necessary, to work with comparatively large single protein crystals —up to 0 2 mm in one dimension The first step in the present crystal analysis of insulin was therefore the preparation of such crystals

THE CRYSTALLIZATION OF INSULIN

A sample of crystalline insulin prepared by Messrs Boots Pure Drug Co Ltd was given for this work by Professor F L Pyman This consisted of very minute rhombohedral crystals and a number of attempts at recrystallization were made in order to grow larger crystals from these Various methods described in the literature were tried but only one recrystallization was successful This followed a slight modification of D A Scott's (1934) later method and the details of the actual experiment are as follows

15 8 mg of insulin were dissolved in 1 c c of water + 1 drop NHCl Meanwhile 12 5 c c of phosphate buffer (Na_2HPO_4 KH_2PO_4 prepared as described by Scott) were diluted with an equal bulk of distilled water and warmed to 55° C To this was added the insulin solution and 2 c c of acetone and the *p*H adjusted to 6 2 by the addition of warm 0 9*N* NH_4OH The solution was set to cool slowly in a cavity in a copper tank filled with water at 55° C At the end of three days the solution was examined and found to contain no crystals The solution was warmed again, a few drops of 0 5 % zinc chloride solution added and the *p*H readjusted to 6 2 There was a slight cloudy precipitate, which mainly dissolved on warming The solution was again allowed to cool slowly for three days as before and at the end of this time comparatively large crystals were found to have formed These were filtered, washed first with the buffer solution, then with distilled water and finally with a little alcohol, and air dried They were very faintly discoloured The main part of the sample was used for a test of the biological activity of the preparation, which was found to be 24 international units per mg This shows that these large crystals are biologically active insulin

---

* It is however possible that the lines found by Clarke and Corrigan indicate that the structure now put forward is a pseudo structure of the real one, or that they were dealing with a polymorphic modification

MORPHOLOGY AND OPTICS OF INSULIN THE QUESTION OF POLYMORPHISM

The crystals prepared as above are shown in the photograph (Pl 9(*a*)) They are essentially very flat rhombohedra, the faces developed being probably those of the simple rhombohedron {100}, or giving hexagonal indices $\{10\bar{1}1\}$ The larger crystals grow unevenly and present the appearance of irregular six lobed stars up to 0 2 mm across and 0 05 mm thick The crystals also frequently grow in pairs united at the ends of their trigonal axes There seems no doubt that the crystalline form here is essentially that found in Abel's (1926) first preparation of insulin crystals and further very well illustrated by Harington and Scott (1929) When viewed along the trigonal axis in convergent light these crystals showed a positive uniaxial figure

Under certain conditions insulin crystallizes in a rather different form, needles or prisms, elongated along the trigonal axis and showing rhombohedral end faces combined with steep trapezohedra (Abel, Geiling, Rouiller, Bell and Wintersteiner 1927 cf Hill and Howitt 1936, p 51) The markedly different appearance of these crystals has suggested polymorphism to a number of observers, and as a polymorphic modification would be particularly valuable in assisting a structure analysis, a preparation of "needle" insulin was examined by the X-ray method Two preparations were available, kindly supplied by Professor F L Pyman The first consisted of very small crystals, roughly wedge-shaped but partly redissolved at the edges, less than 0 05 mm long In appearance and in the frequent occurrence among them of cross-formed twins these crystals seem identical with those described by E B Mathews as present in one of the preparations of Abel and his co-workers (1927, cf Hill and Howitt 1936, p 51) The second preparation examined consisted of much larger roughly needle-shaped masses up to 0 3 mm long which extinguished uniformly between crossed nicols and which although showing no identifiable crystal faces were mainly single crystals (Pl 9(*b*)) The birefringence in both cases was positive The X-ray examination proved that the same crystal structure was present both in these large needle-shaped crystals and in the original flat rhombohedral crystals (Crowfoot 1937)

But although these particular different forms of insulin crystals are not polymorphs, this possibility still remains for certain other preparations of which descriptions appear in the literature Two, particularly, are the crystals having the form of rhombic dodecahedra described by Mathews (Abel and others 1927, cf Hill and Howitt 1936, p 51) and the dumb-bell shaped crystals prepared by Scott (1932) In one recrystallization of insulin, crystals have however been obtained which simulate both these forms and yet seem most probably to be no more than habit variants of the rhombohedral crystals This particular preparation was carried out as follows

10 mg of needle insulin were dissolved in 1 c c of $N/6$ acetic acid and treated with 0 4 c c brucine acetate and 0 2 c c of 13 5% pyridine The main part of the preparation seemed to be precipitated at this stage and the flocculent precipitate was therefore isolated by centrifuging and removing the supernatant liquid It was then dissolved in $N/15$ disodium hydrogen phosphate and $N/6$ acetic acid was added to slight turbidity, as described by Abel On standing in a refrigerator overnight a beautiful crop of very small crystals was obtained

The crystals consisted mainly of a form showing rhombohedral and trapezohedral faces almost equally developed, which gave them at first sight the appearance of rhombic dodecahedra They have, however, positive birefringence and trigonal, certainly not cubic symmetry Amongst these crystals there was a small number showing no rhombohedral faces These were bounded by the basal plane and three trapezohedral faces and were partly redissolved, so that accurate measurement was not possible Many of them were joined end to end, tapering inwards to the join, and they then looked very similar to the "dumb bell" crystals pictured by Scott These crystals also had positive birefringence and, as in the case of the pseudo-rhombic dodecahedra, from this fact and the trigonal symmetry there seems no reason to presume the presence of a different crystal structure The crystals were unfortunately too small for an X-ray test which would be conclusive And it is of course in any case still possible that both Mathews' cubic and Scott's "dumb-bell" crystals are different from either of these

The variation in habit is itself of great interest and worthy of further investigation The small number of experiments so far carried out makes it impossible to be certain which of the varying factors pH, phosphate or acetate buffer, or concentration of the different metallic ions, is responsible for the effects observed The X-ray diffraction effects as described below appear to correspond to a fairly large-scale structure within the protein molecule and it would seem possible to have a certain variation, e g in the metal concentration within the crystals, which might affect the habit without altering the main crystal structure responsible for X-ray diffraction

The X-ray photographs of the needle crystals do differ in one respect from those of the flat rhombohedral crystals All show a very much more marked "amorphous" ring at 4 5 A, a ring which is hardly visible on the photographs of the rhombohedral crystals This ring we have come to consider characteristic of amorphous protein and its presence here may well be due to external imperfection in the crystals But it may also indicate that these crystals contain included amorphous protein, the presence of which has itself affected the crystal habit Of course the occurrence of "amorphous" protein would most probably depend on one or other of the factors mentioned above, pH, metal concentration, different salts present and so on

Certain other properties of the rhombohedral insulin crystals lend some support to these conclusions Optically, as mentioned above, the crystals are uniaxial, the sign of the birefringence positive The refractive indices of the needle crystals were estimated by the oil-immersion method as ϵ slightly $< 1\ 537$ and ω slightly $> 1\ 531$ An independent direct measure of the birefringence gave values between 0 004 and 0 008, most probably about 0 007, but exact measurement was difficult, owing to the irregular shape of the crystals On the large rhombohedra, ω appeared much more nearly equal to 1 537, and it seems probable that this is a real difference and further indication that the needle crystals are less perfect than the rhombohedra

As Abel and others (1927) originally showed, the air-dried rhombohedral crystals contain water of crystallization which they lose when dried over 100° in a vacuum The vacuum-dried crystals are very hygroscopic and rapidly regain moisture on exposure to the air, increasing nearly but not quite to their original weight They then show unimpaired birefringence and give sharp X-ray reflexions identical with those of untreated crystals Check analyses of the water content of different samples of insulin are given in Table I to compare with the crystal densities As the measurement of the density of protein crystals presents certain problems it will be considered next in some detail

THE DENSITY OF INSULIN CRYSTALS

G S Adair and M E Adair (1936) have recently shown that in the case of the crystalline proteins edestin, horse serum albumin and sheep and horse haemoglobin, the density of the crystals varies considerably with the medium in which they are suspended The proteins studied all crystallize with a considerable quantity of water of crystallization and the variation seems most readily explicable as due to specific changes in the salt concentration or proportion of this water The crystallographic estimation of the molecular weight of any protein depends upon the crystal density and here it is the density of the crystals in air as employed for X-ray work that is required But all methods of density determination so far available for use with small crystals do involve their suspension in different liquid media and some care must evidently be exercised to choose media which have little or no specific effect on the crystal density In the case of insulin it might be hoped that such effects would be at a minimum owing to the very small proportion of water of crystallization present

Two values for the density of crystalline insulin have been already recorded The figure given by Dr Eyer of 1 315 (Freudenberg 1932) (method not named) is significantly higher than that first found with the large rhombohedral crystals used for the X ray measurements, namely 1 306 ± 0 003 (Crowfoot 1935) The latter value was measured by flotation under centrifugal force in aqueous solution containing zinc sulphate and zinc chloride The discrepancy might be due either to a medium effect or, as was first thought, to crystallographic imperfections such as inclusions in the large crystals which might have given an apparently too low density for these crystals

A more comprehensive test was therefore carried out using three different samples of crystalline insulin the preparations of needle insulin (large crystals) and unrecrystallized "rhombohedral" insulin (very small crystals) described above, and thirdly the large "rhombohedral" crystals used for the X-ray work

Four sets of solutions were prepared, each consisting of one liquid of density greater than that of the crystals and one of less In outline the method employed has been already described (Bernal and Crowfoot 1934 *b*) and differs slightly in procedure from that used by Adair and Adair, since rather small quantities only of insulin were available The measurements were generally begun by suspending the crystals in a mixture of density about 1 315 in a small glass tube, the suspension then being centrifuged for from 5–15 minutes The low-density solution was then added drop by drop and the rise or fall of the crystals under centrifugal force observed after each addition The density of the liquid was measured at the end of each centrifuge run by a Westphal balance and sinker and the limits at which the crystals just rose or sank in the liquid were recorded When the lower limit had been reached, the high density liquid was in turn added drop by drop in order to pass twice over the equilibrium value of the density With longer centrifuge runs it might have been possible to define more closely the limits, but there seemed to be a certain variation in the crystal density itself over this range, which made such further precision of doubtful value

The liquids used were as follows of which 2 and 4 were chosen following their use by Adair and Adair

1 Zinc sulphate in water ($\rho = 1\ 35$) and distilled water

2 Sodium dihydrogen phosphate in water ($\rho = 1\ 357$) and distilled water

3 *o*-dichlorobenzene ($\rho = 1\ 309$), carbon tetrachloride ($\rho = 1\ 58$) and toluene ($\rho = 0\ 87$)

4 Half saturated ammonium sulphate ($\rho = 1\ 059$) and the same solution saturated with sucrose ($\rho = 1\ 37$)

The results are recorded in Table I

TABLE I DENSITY OF INSULIN CRYSTALS

Solutions	Needles	Small rhombohedra	Large rhombohedra
$ZnSO_4\ H_2O$, *p*H 5 0	1 293 ± 0 003	1 296 ± 0 003	1 300 ± 0 003
$NaH_2PO_4\ H_2O$, *p*H 4 5	1 298 ± 0 003	1 299 ± 0 002	1 301 ± 0 002
o dichlorobenzene toluene and carbon tetrachloride	1 292 ± 0 004 ?	1 297 ± 0 002 ?	1 312 ± 0 004 ?
½ sat $(NH_4)_2SO_4$ + sugar	—	—	1 308 ?
Water content %	5 36	4 93	—

Of these the sugar solution proved completely unsatisfactory for the present method Adair and Adair record a general rise of observed crystal density in sugar solutions and this might be considered to be borne out here by the only figures obtained Actually, under the experimental conditions, these very concentrated solutions seem to form density columns, the crystals lodging at different heights in the liquid but neither completely falling nor rising Such sugar solutions could probably be employed with advantage with a set of sinkers of standard densities but are otherwise useless

The mixture of organic liquids was used as some control on the results obtained in aqueous solutions Observed under the microscope the crystals seemed to remain untouched and birefringent in this immersion medium Actually, in bulk, both the needle crystals and small rhombohedral crystals appeared as cloudy masses and the limits were very difficult to observe, though the results obtained do agree approximately with the densities measured in other solutions It is possible that this opacity is caused by the presence in these preparations of small amounts of non-crystalline protein The large rhombohedral crystals remained clear, and with these definite higher densities were observed than in aqueous solution

The two salt solutions were chosen as solutions containing ions commonly present during the crystallization of insulin, though different in the two examples The phosphate solution in particular appeared to be reliable from the results of Adair and Adair The densities found in both agree in showing a small general variation among the different samples, the needle crystals having the lowest, and the large rhombohedral crystals the highest, density This effect is also observed in the organic liquids and bears out the indications from other sources (X-ray work and refractive index) that the needle crystals are the least perfect of the three preparations It therefore seems most reasonable to adopt the results on the large rhombohedral crystals as most reliable, particularly as these were used for the X-ray work

The observed densities of the rhombohedral crystals considered alone show definite variations with the immersion medium from 1 312 in the organic liquids to 1 301 in the salt solutions These differences are considerably smaller than many of those found with the heavily hydrated protein crystals studied by Adair and Adair and suggest that, though some specific medium effect is present, it is relatively unimportant This idea receives some support from the absence of correlation between crystal size and density among the different insulin preparations studied In the circumstances it is obviously impossible to assign a "true" value to the crystal density, but it is very improbable that any error is introduced by adopting a density of 1 298–1 316 for the purposes of molecular weight determination Crystallographic practice would prefer the higher density as most likely to be correct

THE X-RAY MEASUREMENTS

The first X-ray photographs of insulin were taken at Oxford using throughout copper radiation from a Philips' Metallix research X-ray tube These were given exposure times of 15 hr for each oscillation of 5°, with a plate to crystal distance of 6 cm Four different crystals were used, one rotating about the normal to $(10\bar{1}0)$, one about the normal to $(11\bar{2}0)$ and two about the trigonal axis The rhombohedral character of the symmetry was first tested by single oscillation photographs about these three axes and by oscillation photographs taken at 60° and 120° to each other about the trigonal axis The required relations were found to hold without exception, namely (using hexagonal indices) hkl is absent if

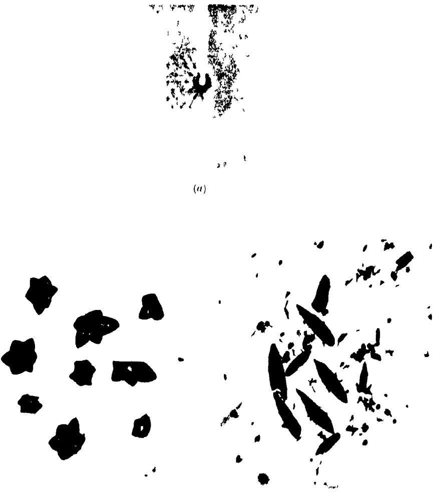
$h-k+l$ is not divisible by three and the symmetry relation $hikl \equiv \bar{i}\bar{k}\bar{h}l$ Using this last identity, a record of the intensities of planes of all the different possible types of indices was obtained by a set of 5° oscillation photographs covering 60° of the total rotation about the trigonal axis This gave in all but a few cases two readings for the intensity of each index type

Through the kindness of Dr Muller, a second set of X-ray photographs of insulin were later taken with the very much more powerful 5 kW tube at the Davy-Faraday laboratory The best photographs were obtained with chromium radiation at a voltage of 28 kV, tube current 150 mA and tube filament 8 amp at 12 V The X-ray tube had a lithium window A fine slit system nearly 6 cm long was employed and the very powerful X-ray beam shielded from the plate by a stop designed by Dr Muller consisting of a lead block 1 mm across at the end of a cylindrical tube 5 mm long which was suspended 2 cm from the crystal Under these conditions it was possible to cut down enormously the exposure times required to obtain diffraction effects The photographs were taken of a single crystal rotating about the normal to $(10\bar{1}0)$ For more accurate measurement of the cell dimensions, two symmetrical photographs were taken, one with the beam from 10° to parallel to the trigonal axis and one at right angles to this For these, the plate to crystal distance was 6 cm To give further readings of the intensities of the X-ray reflexions, a set of 15° oscillation photographs at a plate distance of 4 cm was then made, each having only ¾ hr exposure The photograph reproduced (Pl 9(*c*)), covering the oscillation with the beam 0–10° from parallel to $(1\bar{2}10)$, was, however, given 2 hr and records on a single plate all planes of the type $hik0$ which were observed

Of the needle crystals of insulin, three X-ray photographs only were taken of two different crystals In both cases these were 5° oscillations about the normal to $(10\bar{1}0)$ with the beam direction at the start either parallel or perpendicular to $(1\bar{2}10)$ These were taken at Oxford with copper radiation One photograph was taken of a rhombohedral crystal which had been dried and then allowed to regain water from the air This was also a 5° oscillation about the normal to $(10\bar{1}0)$, beam direction perpendicular to $(1\bar{2}10)$

THE UNIT CELL DIMENSIONS AND MOLECULAR WEIGHT OF INSULIN

From the more accurate photographs taken at the Davy-Faraday laboratory the unit cell dimensions of insulin were recalculated For the



(*a*)

(*b*)

(*c*)

(*a*) Large rhombohedral crystals of insulin (*b*) Large needle crystals of insulin (*c*) X ray photograph of crystalline insulin taken with the rotation axis perpendicular to $(10\bar{1}0)$ and beam direction parallel to $(1\bar{2}10)$

hexagonal cell, $a = 74·8$ A, $c = 30·9$ A, for the rhombohedral unit, $a = 44·4$ A, $\alpha = 114° 48'$ correct to about 0·5 %. There are no systematic series of absent reflexions, but those required by the rhombohedral symmetry and, as the photograph reproduced illustrates, no planes of symmetry. The space group is thus $R3$.

The cell molecular weight calculated from the above figures and a density of 1·312 is 39,700. It may be as high as 40,800 or less probably as low as 38,700. The actual molecular weight of protein in the cell depends on the water of crystallization present. Abel's original measurements on drying the crystals at 104° under reduced pressure gave an average value of 5·35 %, and this is borne out by estimations on certain of the crystals used (p. 588). This protein molecular weight may therefore be calculated as 37,600 (limits 38,900, 36,000).

It is clear from these measurements that the crystal unit cell of insulin contains one "molecule" on the Svedberg scale, but it is interesting to compare the figures obtained above more exactly with those found by the ultracentrifuge method. The first determination by Svedberg and Sjogren (Svedberg 1931, Sjogren and Svedberg 1931) gave 35,100 as the molecular weight of insulin, but it is probable that here, as in most of the earlier experiments with the ultracentrifuge, true sedimentation equilibrium was not obtained. A later re-estimation (unpublished work by A. G. Polson (Svedberg 1937 *b*) gives a value of 40,900, which is somewhat higher than the calculated weight of protein in the crystal unit cell of insulin, 37,600. It is doubtful whether this difference between the X-ray and ultracentrifuge results is outside the limits of the experimental error of both series of measurements, and the same may probably be said with reference to the shape of the "molecule", which is spherical according to Svedberg's determination, while the unit cell is a markedly oblate spheroid. It may be that this last deviation is due to further hydration of the protein in solution, since it is already partly hydrated in the crystal (cf. the calculation by Adair and Adair on haemoglobin). Such hydration would also alter the density of the moving particle which is introduced in Svedberg's equations, but this should not necessarily affect the deduction of the actual protein molecular weight. The corrections to be applied in such a molecular weight measurement either for a change of density or for the non-spherical nature of the falling particle seem definitely smaller than the present experimental error.

It should be pointed out that while the mass of protein in the unit cell is equivalent to that of one Svedberg unit, the X-ray data do not actually supply any evidence that this mass constitutes a single molecule in the crystal The same relation to the Svedberg unit would be obtained by the presence in the crystal unit of $3n$ chemical molecules which subsequently combined in solution to form a single aggregate

THE INTENSITIES OF THE X-RAY REFLEXIONS

The relative intensities of the X-ray reflexions were estimated by eye and assigned values on an arbitrary numerical scale The scale intervals were checked by comparisons between the α and β reflexions from different planes and also against a plate on which one X-ray spot had been exposed for different measured lengths of time These estimated intensities were corrected by the Lorentz and polarization factors and also, when not occurring on the equatorial line, according to the relation found by Cox and Shaw (1930), in order to give a set of numerical readings for each plane, F'^2 proportional to F^2 Both the set of photographs—oscillations about perpendicular to $(10\bar{1}0)$—taken at the Davy-Faraday laboratory and the oscillation photographs about [0001] taken at Oxford were used in preparation of the values finally adopted In the first case comparison between different plates was made mainly by assuming equal exposure, generally checked by overlapping spectra In the second case many of the exposures were obviously of different value and here the comparison was based mainly on the occurrence on every film of some reflexion of the type *hiko* As all *hiko* reflexions occurred together on another plate their interrelation was easily found In all but two cases of weak reflexions (missed in one series or the other through a small degree of missetting) there were at least two occurrences of every index type and generally more Good agreement was found between the different values of F'^2 assigned to the same reflexion occurring on different photographs and also on different layer lines of the same photograph The last, particularly, is some evidence that the relative scale of intensities is fairly correct The values finally adopted are listed in Table II

PATTERSON-HARKER ANALYSIS OF INSULIN

The set of intensities of X-ray reflexions from insulin recorded above have certain characteristics which must be taken into account in their analysis In the first place they are a very limited series—only 59 terms in all No reflexions have been observed from planes with a spacing smaller than 7 05 A, and at this distance the intensity is extremely weak There seems to be a definite termination of the series in this region There should therefore, for syntheses formed with these terms, be none of the errors ordinarily introduced by arbitrarily cutting short the series On the other hand it is impossible to expect to obtain information from this set of intensities about the most intimate structure of insulin—actual distances between neighbouring atoms The X-ray reflexions observed must correspond to distances between scattering masses of a fairly large-scale structure within the molecule

TABLE II—RELATIVE INTENSITIES OF X-RAY REFLEXIONS FROM INSULIN

F'^2 values (proportional to F^2)

Indices	F'^2	Indices	F'^2	Indices	F'^2	Indices	F'^2
$30\bar{3}0$	50	$3\bar{1}\bar{2}1$	19	$10\bar{1}\bar{2}$	3	$7\bar{3}\bar{4}\bar{2}$	7
$60\bar{6}0$	7	$21\bar{3}1$	4	$20\bar{2}2$	5	$43\bar{7}\bar{2}$	4
$11\bar{2}0$	10	$4\bar{1}\bar{3}\bar{1}$	12	$40\bar{4}\bar{2}$	—	$8\bar{3}\bar{5}2$	—
$22\bar{4}0$	8	$31\bar{4}\bar{1}$	4	$50\bar{5}2$	4	$53\bar{8}2$	—
$33\bar{6}0$	8	$6\bar{1}\bar{5}1$	3	$70\bar{7}\bar{2}$	—		
$44\bar{8}0$	9	$51\bar{6}1$	4	$31\bar{2}\bar{2}$	11		
$51\bar{4}0$	12	$7\bar{1}\bar{6}\bar{1}$	11	$21\bar{3}\bar{2}$	9	0003	3
$41\bar{5}0$	*t*–	$6\bar{1}7\bar{1}$	—	$4\bar{1}\bar{3}2$	3	$30\bar{3}3$	3
$7\bar{2}\bar{5}0$	11	$5\bar{2}\bar{3}1$	3	$31\bar{4}2$	30	$30\bar{3}\bar{3}$	1
$52\bar{7}0$	12	$32\bar{5}1$	4	$6\bar{1}\bar{5}\bar{2}$	9	$22\bar{4}3$	4
$8\bar{1}\bar{7}0$	22	$6\bar{2}\bar{4}\bar{1}$	6	$51\bar{6}\bar{2}$	4	$5\bar{1}\bar{4}3$	3
$71\bar{8}0$	6	$42\bar{6}\bar{1}$	25	$7\bar{1}\bar{6}2$	4	$51\bar{4}\bar{3}$	2
$9\bar{3}\bar{6}0$	3	$8\bar{2}\bar{6}1$	8	$61\bar{7}2$	2	$41\bar{5}3$	—
$63\bar{9}0$	*t*–	$62\bar{8}1$	3	$5\bar{2}\bar{3}\bar{2}$	3	$41\bar{5}\bar{3}$	3
		$9\bar{2}\bar{7}\bar{1}$		$32\bar{5}\bar{2}$	6		
$10\bar{1}1$	36	$72\bar{9}\bar{1}$	—	$6\bar{2}\bar{4}2$	—		
$20\bar{2}\bar{1}$	3	$7\bar{3}\bar{4}1$	5	$42\bar{6}2$	4		
$40\bar{4}1$	7	$43\bar{7}1$	6	$8\bar{2}\bar{6}\bar{2}$	—		
$50\bar{5}\bar{1}$	4	$8\bar{3}\bar{5}\bar{1}$	7	$62\bar{8}\bar{2}$	—		
$70\bar{7}1$	10	$53\bar{8}\bar{1}$	2	$9\bar{2}\bar{7}2$	3		
		$9\bar{4}\bar{5}1$	—	$72\bar{9}2$	—		
		$54\bar{9}1$	3				

To obtain a picture of all the information that can be derived, without any working hypotheses directly from these X-ray reflexions, a series of Patterson-Fourier syntheses was prepared using the arbitrary F'^2 values recorded above as coefficients of the terms In these two-dimensional diagrams the distances of the various peaks from the origin measure in length and direction prominent interatomic distances in the structure The first diagram, fig 1, was formed using only the intensities of the *hiko* reflexions This gives a projection of the whole Patterson structure on the (0001) plane Owing to the presence of the rhombohedral lattice the primary pattern repeats at ⅓, ⅔ and ⅔, ⅓ The peaks at these positions correspond to distances between atoms at $z = \frac{1}{3}$ and $z = \frac{2}{3}$ respectively and have necessarily the same value as the origin peak, which states that every atom is at zero distance from itself The remaining diagrams illustrated are essentially sections cut through the structure along different lines To form these the F'^2 values of all $(hikl)$ planes were used, as first described by

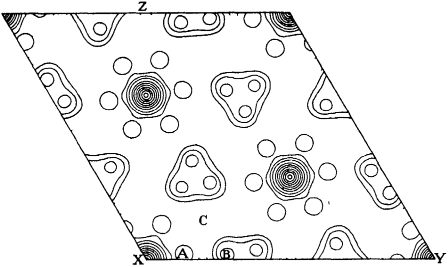


FIG 1 $P(xy)$ for insulin derived from Patterson Fourier analysis Contours at 5 units apart

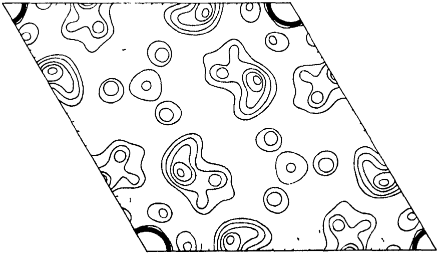


FIG 2 $P(xyo)$ for insulin derived from Patterson Harker analysis Contours at 50 units apart, those above 250 being omitted (and similarly in figs 3, 4 and 5)

D Harker (1936) The second diagram, fig 2, is a section cut parallel to (0001) at the level $z = 0$ in the Patterson structure and is obtained by the relation

$$P(xyo) = \Sigma_h \Sigma_k \cos(hx + ky) \ \ (\Sigma_l |F(hkl)|^2)$$

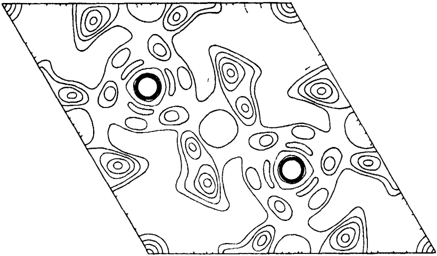


FIG 3 $P(xy\frac{1}{2})$ for insulin derived from Patterson Harker analysis

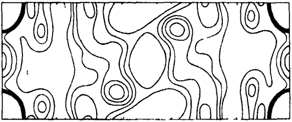


FIG 4 Patterson Harker analysis of insulin Section in fig 1 at XY

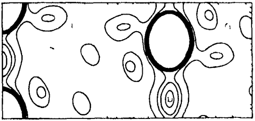


FIG 5 Patterson Harker analysis of insulin Section in fig 1 at XZ

Fig 3 is cut parallel to (0001) at a height $z = \frac{1}{2}$ and the remaining two diagrams, figs 4 and 5, are cut perpendicular to the rest along the lines XY and XZ of projection 1 The summation of the Fourier series involved was carried out by the method of Beevers and Lipson (1936*a*, *b*) using the strips supplied by them As it was evident from the number of terms involved that no very fine structure could be found, the points in the projections were calculated at intervals of $\frac{1}{30}a$ (c 2 5 A) and $\frac{1}{15}c$ (2 A) The contour lines are drawn at intervals of 50 units in projection 1 and 500 units in the remaining projections, those in negative regions being dotted As in most cases the slopes involved are very gradual, the choice of the exact areas bounded by the contour lines is somewhat arbitrary and these should not be given a too rigid interpretation

Between them these five diagrams present a fairly complete survey of what may be called—as above—the Patterson structure of insulin Owing to the rhombohedral symmetry, the diagrams, figs 2 and 3, sections cut parallel to (0001) at 0 and $\frac{1}{2}$, actually supply sections of the structure in this direction at intervals of $z = \frac{1}{6}$, the distribution at the point $\frac{2}{3}$, $\frac{1}{3}$, in fig 2, being, for example, the same as that at 0 at a height of $z = \frac{1}{3}$, and so on Some check on the interreliability of the intensity data can be obtained by addition of the density values derived from points on these two projections 2 and 3 at z intervals of $\frac{1}{6}$ and then comparing the picture formed with projection 1, since projection 1 is derived from only *hiko* reflexions and the others from all *hikl* reflexions The agreement is fairly good, and even better over the lines AB and AC, where density values have been calculated at intervals of $z = \frac{1}{15}$

The peak structure shown by the five projections seems comparatively simple There are two main series of peaks (A and B), one at a distance of 10 A from the origin parallel to (0001) and the other at about 22 A measured similarly These peaks have considerable extension at right angles to (0001), the main intensity maxima appearing some distance above and below the height $z = 0$ There is further a marked peak at $x = 0$, $y = 0$, $z = 0$ 5 (D), and a considerably weaker series, of which defining co-ordinates may be given as $x = 0$ 29, $y = 0$ 19, $z = 0$ (C) Owing to the poor resolution and overlapping it is impossible to pretend to any accuracy in assigning positions to the different peaks, and it seems probable that to a large degree their irregular outlines are due to their composite character and should be taken into account in their interpretation Approximately, however, the co-ordinates of the main peaks may be recorded as in Table III, which gives one selection of peaks which by repetition according to rhombohedral symmetry would produce the whole Patterson structure It is of course possible that some other selection might be made which would throw more light on the real molecular structure

In order to pass from the Patterson structure to the molecular structure it is necessary to reduce the peak system shown in the Patterson diagrams to its simplest terms Certain interatomic distances which appear must be derived from others in real space The patterns found are comparable in simplicity with those obtained in molecular structures of quite small dimensions, e g those of pentaerythritol (Llewellyn, Cox and Goodwin 1937), where the peaks represent distances between single atoms At first sight they suggest that here we might be dealing with masses about 10 A, across which simulate the properties of single atoms in simpler structures But even assuming the presence of such masses it has not so far proved possible to offer a derivation of the Patterson diagrams in terms of an actual molecular arrangement And the problem may, of course, be much more complicated, each peak being the result of a comparatively small proportion of scattered interatomic distances Some measure of the absolute value of the intensities of the X-ray reflexions, which will next be undertaken, may offer an easier approach to a solution

TABLE III—APPROXIMATE CO-ORDINATES OF PEAKS IN PATTERSON SERIES

	x	y	z	d in A
A	0 147	0 027	0 12	10 5
			↑	↑
			0	9 5
			↓	↓
	0 127	0 113	0 25	12 5
B	0 313	0 047	0 17	22 5
			↑	↑
			0	21 7
			↓	↓
	0 30	0 047	0 33	24 0
C	0 29	0 19	0	19 5
D	0	0	0 5	15 0

GENERAL CONCLUSIONS THE PATTERSON ANALYSIS OF INSULIN AND PROTEIN STRUCTURE

Even without the derivation of a molecular structure for insulin from the Patterson analysis, certain conclusions relevant to the main problem of protein structure may be drawn The first question raised at the beginning of this paper was that of a check on existing theoretical speculations It must be said that the patterns calculated do not appear to have any direct relation either to the cyclol or to the various chain structures put forward for the globular proteins From the cyclol structure (Wrinch 1937 *b*) one would expect many more peaks in the Patterson synthesis than do actually occur For example, besides peaks at about 10 A parallel to

(0001) there should also be peaks at 5 and 15 A It is less easy to control the folded-chain hypotheses, as for these no exact models have yet been calculated, but in outline the picture does not suggest a chain structure There is no indication in the diagrams of layers parallel to (0001) as some theories would require, and the idea of a spiral chain which has otherwise some attraction does not, at least in its general form, account for the definite distribution of 10 and 20 A interatomic distances in separate peaks It is naturally difficult to rely on these conclusions without still more exact theories involving the actual positions of the different amino-acid residues in the protein molecule But the general lack of positive agreement remains evident

Perhaps the most unexpected information derived from the preliminary data on insulin was, as mentioned above, that the unit of weight about 39,000 had apparently trigonal symmetry This may be considered to receive some support from the chemical analyses of several proteins by M Bergmann and C Niemann (1937), who find that the amino-acid residues constantly recur in multiples of three The most natural explanation of trigonal symmetry is that the Svedberg unit is actually an aggregate of $3n$ chemical molecules It is *a priori* less likely that a molecule of such magnitude should possess real trigonal symmetry, though it should perhaps be pointed out that this is actually developed by the cyclol theory It is, however, doubtful whether the facts justify pushing this argument further The new experimental evidence is still unable to exclude the possibility that the trigonal symmetry shown by the crystal and the diffraction effects is due to statistical twinning and hence is not a property of the Svedberg "molecule" itself And even apart from this reservation, the present examination shows that a somewhat less rigid interpretation of the facts is possible It still seems necessary to postulate an approximately trigonal distribution of masses within the Svedberg unit, but this distribution need not extend to atomic dimensions Irregularities of the order of 1 or 2 A would plainly not be indicated by the observed diffraction effects It was hoped that more powerful X-ray apparatus would reveal a finer structure, but such if present did not appear on the exposures taken with the Davy-Faraday 5 kW tube And it is at least plausible that the very definite cutting short of the series of reflexions at spacings from 8–7 A is a positive indication that there are irregularities within the structure below this limit (see below)

The most interesting fact that has emerged from the Patterson analysis is the very definite distribution in space of groups of interatomic distances above this limit of 7 A, particularly those of approximately 10–12 A and

22 A The appearance of these sets of distances immediately relates the internal structure we are examining in insulin with the diffraction effects previously observed in other proteins, both fibrous and crystalline We may connect these 10–11 A and 22 A distances with the spacings commonly observed in denatured protein crystals such as pepsin and edestin (Astbury 1934) and even proteins such as gelatin (Gerngross, Herrmann and Abitz 1930) The 10–11 A distance then appears identical with Astbury's "side chain" spacing in keratin or myosin And this is supported by a striking relationship between these X-ray results for insulin and those on excelsin According to the preliminary measurements of Astbury, Dickinson and Bailey (1935), excelsin crystallizes in a rhombohedral structure with one molecule in the unit cell, not unlike that of insulin, though the scale is much greater The crystals of excelsin examined had, however, been kept for some time, which resulted in partial degeneration, and the diffraction patterns obtained showed, besides true crystal X-ray reflexions, an oriented fibre pattern With the X-ray beam perpendicular to the basal plane this took the form of six blurred sickles with spacings of 11 4 A around the trigonal axis, while at right angles the fibre reflexion had a spacing of roughly 4 5 A The orientation of the 11 4 A spacing of the excelsin fibre pattern parallel to the crystal basal plane is the same as that of the 10–11 interatomic distances in insulin And these consequently supply further evidence of the close relationship between the actual intermolecular structures of the fibre and globular proteins It is difficult not to connect these spacings also with the intramolecular patterns obtained from the virus proteins From these Bernal and Fankuchen (1937) have been led to deduce the presence of submolecules 20 × 20 × 22 A, which are again divided into nearly identical units of half these dimensions But whether these spacings of 10 and 20 A, which occur in all these different types of proteins, are actually due to distances between chains held apart by their side groups, as Astbury advocates, remains still to be proved The evidence of the ultracentrifuge measurements, which show a reversible dissociation of insulin outside the *p*H stability range 4 5–7, may perhaps be quoted as some support for an alternative hypothesis that they are due to definite massive subdivisions within the protein molecule Whatever their explanation, the Patterson analysis of insulin provides a first exact description of their vectorial distribution in space

The identification of these 10 A spacings in insulin raises the problem of the occurrence of the second generally found protein spacing of 4 5 A The crystal diffraction effects observed in insulin which end at 7 A can give no information as to the origin of the 4 5 A ring found on the powder photographs Yet this ring must be due to the internal structure of the protein The intensity of reflexion at any angle depends upon the product of what may be called the atomic and molecular scattering factors at that angle And accordingly with a perfectly ordered protein crystal one might expect recurrence of sharp diffraction effects in the region of 4 5 A, which would be of particular importance for a study of the most intimate structure of the protein molecule Their absence in insulin suggests on this reasoning some actual disorientation of atoms within the molecule, rather than mutual disorientation of the protein molecules But either effect might be caused by the processes involved in the preparation of the crystals and their present physical condition And in this case it is reasonable to hope for results on other protein crystals which will carry the X-ray analysis deeper into the problem of their structure

ACKNOWLEDGEMENTS

My first thanks are due to Professor R Robinson, without whose interest the present work would not have been begun, and also to Professor F L Pyman for the gift of the insulin used in this research For the tests on the biological activity of the recrystallized insulin I am indebted to Sir Henry Dale and Mr E C Marks, and for facilities during the summer of 1936 at the Davy-Faraday laboratory to Sir William Bragg and the Managers of the Royal Institution I should also like to thank particularly Dr Müller and Mr Smith for invaluable assistance in carrying out the experiments at the Davy-Faraday laboratory, and Mr J D Bernal for much help and advice in the preparation of this paper

SUMMARY

X-ray crystallographic measurements show that insulin crystallizes in the rhombohedral system, space group $R3$ The dimensions of the unit cell are $a = 44\ 4$A, $\alpha = 114° 48'$ or, referred to hexagonal axes, $a = 74\ 8$, $c = 30\ 9$A correct to about 0 5 % The densities of three samples of crystalline insulin have been measured in four different immersion media and show small variations between values of 1 292 and 1 316 From these and the cell dimensions the unit cell molecular weight may be calculated as most probably 39,700 (maximum 40,800, minimum 38,700) The crystals contain water which they lose when dried at 104° in a vacuum and this has been estimated at 5 35 % The molecular weight of protein in the unit cell is hence calculated as 37,600 (38,900–36,000), which is of the order of magni-

tude though rather less than the molecular weight of insulin found by Svedberg There is, however, no crystallographic evidence that this weight constitutes a single molecule

The intensities of all the X-ray reflexions that could be observed from air-dried single crystals of insulin have been estimated visually and are recorded as terms proportional to F^2 From these terms a series of Patterson-Harker Fourier syntheses has been formed which shows the distribution of the main interatomic distances within the insulin molecule The patterns correspond to a fairly large scale structure and as such cannot give information on the placing of individual atoms or as to whether the trigonal symmetry shown is necessarily carried to atomic dimensions The most important interatomic distances that appear are at 10 and 22 A roughly parallel to the basal plane It is suggested that these are related to the characteristic X-ray spacings of many proteins both globular and fibrous, and the virus proteins

REFERENCES

Abel, J J 1926 *Proc Nat Acad Sci* **12**, 132

Abel, J J , Geiling, E M K , Rouiller, C A Bell, F K and Wintersteiner, O 1927 *J Pharmacol* **31**, 65

Adair, G S and Adair, M E 1936 *Proc Roy Soc* B, **120**, 422

Astbury, W T 1934 *Kolloidzschr* **69**, 340

Astbury, W T , Dickinson, S and Bailey, K 1935 *Biochem J* **29** 2351

Astbury, W T and Marwick, T C 1932 *Nature, Lond* , **130**, 309

Astbury, W T and Sisson, W A 1935 *Proc Roy Soc* A, **150**, 533

Astbury, W T and Woods, H J 1933 *Phil Trans* A, **232**, 333

Beevers, C A and Lipson, H 1936*a* *Proc Phys Soc* **48**, 772

— — 1936*b* *Nature, Lond* , **137**, 825

Bergmann, M and Niemann, C 1937 *J Biol Chem* **118**, 301

Bernal, J D and Crowfoot, D 1934*a* *Nature, Lond* , **133**, 794

— — 1934*b* *Nature, Lond* , **134**, 809

Bernal, J D and Fankuchen, I 1937 *Nature, Lond* , **139** 923

Clarke, G L and Corrigan, K E 1932 *Phys Rev* ii, **40**, 639

Cox, E G and Shaw, W F B 1930 *Proc Roy Soc* A, **127**, 71

Crowfoot, D 1935 *Nature, Lond* , **135**, 591

— 1937 *Nature, Lond* , **140**, 149

Fankuchen, I 1934 *J Amer Chem Soc* **56**, 2398

Freudenberg, K 1932 *Hoppe Seyl Z* **204**, 233

Freudenberg, K , Discherl, W and Eyer, H 1930 *Hoppe Seyl Z* **187**, 89

George, W H 1928 *Proc Leeds Phil Lit Soc* **1**, 412

Gerngross, O , Herrmann, K and Abitz, W 1930 *Biochem Z* **228**, 414

Harington, C R and Scott, D A 1929 *Biochem J* **23**, 390

Harker, D 1936 *J Chem Phys* **4**, 381

Hill, D and Howitt, F 1936 "Insulin", p 51 Hutchinson

Llewellyn, F J , Cox, E G and Goodwin, T H 1937 *J Chem Soc* p 883

Magnus Levy, A , Meyer, K H and Lotmer, W 1936 *Nature, Lond* , **137**, 616
Patterson, A L 1934 *Phys Rev* **46**, 372
— 1935 *Z Kristallogr* **90**, 517, 543
Pauling, L and Mirsky, A E 1936 *Proc Nat Acad Sci* **22**, 439
Robertson, J M 1934 *Proc Roy Soc* A, **146**, 473
Scott, D A 1932 *Trans Roy Soc Can* [iii], **26**, v, 275
— 1934 *Biochem J* **28**, 1596
Sims, H and Scott, D A 1930 *Trans Roy Soc Can* [iii], **24**, v, 117
Sjögren, B and Svedberg, T 1931 *J Amer Chem Soc* **53**, 2657
Svedberg, T 1931 *Nature, Lond* , **127**, 438
— 1937*a* *Chem Rev* **20**, 81
— 1937*b* *Nature, Lond* , **139**, 1060
Wrinch, D M 1937*a* *Proc Roy Soc* A, **160**, 59
— 1937*b* *Science*, **85**, 566
Wyckoff, R W G and Corey, R B 1935 *Science*, **81**, 365
— — 1936 *J Biol Chem* **114** 407

INDEX TO VOLUME CLXIV (A)

Adsorption of argon, nitrogen and oxygen (Wilkins), 510
Adsorption of gases at solid surfaces (Wilkins), 496
Aerofoils, two dimensional hydrodynamical theory (Morris), 346
Appleton (E V) and Piddington (J H) The reflexion coefficients of ionospheric regions, 467

Berg (W F) Crystal growth from solutions, 79
Best (J E), Farmer (F T) and Ratcliffe (J A) Studies of region E of the ionosphere, 96
Bhabha (H J) On the penetrating component of cosmic radiation, 257
Bjerge (T) The production of neutrons by bombardment of beryllium with α particles, 243
Black (A N) and Southwell (R V) Relaxation methods applied to engineering problems II Basic theory, with applications to surveying and to electrical networks, and an extension to gyrostatic systems, 447
Blackman (M) On anomalous vibrational spectra, 62
Bleaching of visual purple solutions (Dartnall, Goodeve and Lythgoe), 216

Cameron (W H B) *See* Elliott and Cameron
Clews (C J B) and Schoszberger (F) Structure of stretched rubber, 491
Cosmic radiation (Bhabha), 257
Coulson (C A) The electronic structure of some polyenes and aromatic molecules IV The nature of the links of certain free radicals, 383
Crystal growth (Berg), 79
Cundy (H M) The rotational energy levels of a diatomic molecule in a tetrahedral field, 420

Dartnall (H J A), Goodeve (C F) and Lythgoe (R J) The effect of temperature on the photochemical bleaching of visual purple solutions, 216
Dibenzyl series, X ray analysis V (Robertson and Woodward), 436

Electronic structure of some polyenes and aromatic molecules IV, V, VI (Coulson, Wheland, Penney and Kynch), 383, 397, 409
Elliott (A) and Cameron (W H B) The emission band spectrum of chlorine (Cl_2^+) II, 531

Farmer (F T) *See* Best, Farmer and Ratcliffe

Galvano magnetic effects in bismuth alloys (Thompson), 24
Goldsbrough (G R) The seiches in a strait connecting two seas, 1
Goodeve (C F) *See* Dartnall, Goodeve and Lythgoe
Goodwin (T H) and Hardy (R) The crystal structure of pentaerythritol tetracetate, 369
Gurney (R W) and Mott (N F) The theory of the photolysis of silver bromide and the photographic latent image, 151

Hardy (R) *See* Goodwin and Hardy
Hartree (D R) Self consistent field with exchange for calcium, 167
Howarth (L) On the solution of the laminar boundary layer equations, 547
Howarth (L) *See* Kármán and Howarth
Hydrogen, new spectrum bands (Richardson), 316
Hydrolysis of methyl halides (Moelwyn-Hughes), 295
Hyperfine structure of aluminium (Jackson and Kuhn), 48

Instability of heated fluids (Krishna Chandra), 231
Ionosphere, region *E* (Best, Farmer and Ratcliffe), 96
Ionospheric regions, reflexion coefficients (Appleton and Piddington), 467

Jackson (D A) and Kuhn (H) Hyperfine structure and nuclear moments of aluminium, 48
Jahn (H) Stability of polyatomic molecules in degenerate electronic states II Spin degeneracy, 117
Jeffreys (H) Significance tests for continuous departures from suggested distributions of chance, 307

Kármán (T de) and Howarth (L) On the statistical theory of isotropic turbulence, 192
Krishna Chandra Instability of fluids heated from below, 231
Kuhn (H) *See* Jackson and Kuhn
Kynch (G J) *See* Penney and Kynch

Laminar boundary layer equations (Howarth), 547
Lythgoe (R J) *See* Dartnall, Goodeve and Lythgoe

Moelwyn Hughes (E A) The hydrolysis of the methyl halides, 295
Morris (Rosa M) The two dimensional hydrodynamical theory of moving aerofoils II, 346
Mott (N F) *See* Gurney and Mott

Penney (W G) and Kynch (G J) The electronic structure of some polyenes and aromatic molecules VI Phenylethylene, stilbene, tolane and the phenylmethyl radical, 409
Pentaerythritol tetracetate (Goodwin and Hardy), 369
Photolysis of silver bromide (Gurney and Mott), 151
Piddington (J H) *See* Appleton and Piddington
Progressive lightning IV (Schonland), 132

Ratcliffe (J A) *See* Best, Farmer and Ratcliffe
Relaxation methods applied to engineering problems II (Black and Southwell), 447
Richardson (O W) New bands ending on the $1s\sigma 2p\sigma\, {}^1\Sigma_u$ state of H_2, 316
Robertson (J M) and Woodward (I) X ray analysis of the dibenzyl series V Tolane and the triple bond, 436
Rotational energy levels of a diatomic molecule in a tetrahedral field (Cundy), 420
Rubber, stretched, structure (Clews and Schoszberger), 491

Schonland (B F J) Progressive lightning IV The discharge mechanism, 132
Schoszberger (F) *See* Clews and Schoszberger
Seiches in a strait (Goldsbrough), 1
Self-consistent field with exchange for calcium (Hartree), 167
Significance tests (Jeffreys), 307
Southwell (R V) *See* Black and Southwell
Spectra, anomalous vibrational (Blackman), 62
Spectrum, new bands of H_2 (Richardson), 316
Spectrum of chlorine (Cl_2^+) II (Elliott and Cameron), 531
Spectrum of turbulence (Taylor), 476
Stability of polyatomic molecules (Jahn), 117

Taylor (G I) Production and dissipation of vorticity in a turbulent fluid, 15
Taylor (G I) The spectrum of turbulence, 476
Thompson (N) Galvano magnetic effects in bismuth alloys, 24
Turbulence, spectrum (Taylor), 476
Turbulence, theory (Kármán and Howarth), 192
Turbulent fluid, vorticity (Taylor), 15

Vorticity in a turbulent fluid (Taylor), 15

Wheland (G W) The electronic structure of some polyenes and aromatic molecules V A comparison of molecular orbital and valence bond methods, 397
Wilkins (F J) Statistical mechanics of the adsorption of gases at solid surfaces, 496
Wilkins (F J) The adsorption of argon, nitrogen and oxygen on smooth platinum foil at low temperatures and pressures, 510
Woodward (I) *See* Robertson and Woodward

END OF THE ONE HUNDRED AND SIXTY FOURTH VOLUME (SERIES A)

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form These abstracts will be issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings"

NOTE—The pagination of the 'Abstracts" will be consecutive through the year and independent of the 'Proceedings" To avoid any confusion with the 'Proceedings", the page numbers will be preceded by the letter S (*i e* Summary)

---

21 JANUARY 1938

Chemical structure in relation to oestrogenic activity Compounds without the phenanthrene nucleus By E C DODDS and W LAWSON

(*Communicated by Sir Henry Dale, F R S -Received* 11 *December* 1937)

The oestrogenic activity of calciferol and clupanodonic acid appeared to indicate that the phenanthrene nucleus was not necessary for oestrogenic activity This view was substantiated by demonstrating the potency of various derivatives of acenaphthene

A large number of other substances, such as carbinols, derivatives of diphenyl methane, diphenyl ethane and diphenyl ethylene have been shown to be oestrogenic Compounds as simple as *p* hydroxy propyl benzene and *p* hydroxy propenyl benzene were found to possess full oestrogenic activity The most potent substances, however, were found among the diphenyl ethylene derivatives

The possibility of the formation of polymerides from *p* hydroxy propenyl benzene was considered

Quantitative spectrographic analysis of biological material III.

By J S Foster, F R S, G O Langstroth and D R McRae (*Received 14 December* 1937)

An internal standard method of quantitative spectrographic analysis, as applied in the determination of Na and K in glandular secretions, is described The probable error in a determination is about 4%

By adding an appropriate foreign substance to the sample in such quantities that the character of the condensed a c spark discharge is determined by this substance, it is possible to use the same working curve for samples of considerably different composition

The sample and internal standard may be separately placed on the electrode by a pipette specially designed for accurate delivery of small volumes The volume of a sample required for the determination of both elements is thus reduced to 0 01 c c

By photographing the spectrum of the sample with an antimony absorption step weakener before the spectrograph slit it is possible to obtain determinations for two or more elements from a single exposure

---

The secretion of protein material in the parasympathetic submaxillary saliva

By G O Langstroth, D R McRae and G W Stavraky (*Communicated by J S Foster, F R S —Received* 14 *December* 1937)

A quantitative study has been made of the variations in composition which occur in the submaxillary saliva of the cat when the intensity of chorda tympani stimulation is altered The relative concentrations of protein material were determined from the absorption spectrum of the saliva, while Na and K concentrations were determined from the emission spectrum Less extensive analyses for other constituents were made chemically

A mathematical theory of the secretion of protein material has been developed It presents the following picture of the secretory processes (1) During stimulation some activating substance is liberated within the gland, (2) this substance sets in operation the water secretion mechanism, controls the membrane permeability, and sets up a reaction, or chain of reactions, resulting in the transformation of granule material to a form readily carried out of the gland by the flow of water It leads to the following expression for the relative concentrations of protein material in two samples (n, i) taken from a series in which the intensity of the stimulus is varied

$$C_i/C_n = D_i/D_n \, e^{-\sum_{n}^{i-1} kB_r}$$

k is a constant, D_i represents the concentration of (Na + K) in the ith sample in milliequivalents per litre, and B_r represents the amount of (Na + K) secreted in the rth sample in milliequivalents The expression represents the observed results remarkably well The theory is also in accord with certain other observed features of the secretory processes

The theory of pressure-ionization and its applications By D S KOTHARI (*Communicated by M N Saha, F R S —Received* 14 *December* 1937)

The main results which have been obtained from the application of the theory of pressure ionization may be summarized as follows

(i) The theory predicts that the stellar material in the interior of the white dwarf stars should be fully ionized

(ii) It predicts that there cannot be a "cold" body (planet of white dwarf) larger in size than Jupiter

(iii) The theory shows that the two heaviest planets (Jupiter and Saturn) have cores composed of metallic hydrogen The terrestrial planets have cores of much heavier metal possibly iron

Comparison of wave-functions for HeH^{++} and HeH^{+} By C A COULSON and W E DUNCANSON (*Communicated by E N da C Andrade, F R S —Received* 14 *December* 1937)

As many different approximations as possible have been used in a comparative study of the wave functions of HeH^{++} and HeH^{+} In the latter ion, the molecular orbital approximation is found to be much better than the electron pair approximation Ionic terms and polar terms are included in the final wave function, and an expansion in terms of spheroidal co ordinates is given The lowest value obtained for the energy of HeH^{+} is -2 935 A, with an internuclear distance of 1 446 A = 0 764 A° The dissociation energy certainly lies within 0 61 and 2 10 V, being probably about 1 5 V The fundamental vibration frequency is 3380 cm^{-1}

The constitution of heavy water By I RAMAKRISHNA RAO and P KOTESWARAM (*Communicated by O W Richardson, F R S —Received* 14 *December* 1937)

The Raman band of heavy water is found to be similar to that of ordinary water both in diffuseness and extent The variations with temperature in the structure of this band are also similar to those of ordinary water These changes are attributed to changes in the relative proportions of (D_2O), $(D_2O)_2$ and $(D_2O)_3$ molecules which are assumed to be in thermal equilibrium By an analysis of the intensity curves of the band, the proportions of the three polymers at different temperatures are calculated The results are compared with those of water and the points of similarity and differences explained on the basis of their other physical properties

The production of γ-rays by neutrons By E H S BURHOP, R D HILL and A A TOWNSEND (*Communicated by T H Laby, F R S —Received* 14 *December* 1937)

Measurements have been made of the absorption in boron of those neutrons which excite γ rays in specimens of cadmium, silver (of two thicknesses), arsenic, antimony, iodine and mercury The measurements were carried out using thermal neutrons in the case of cadmium and silver (thin specimen) and non thermal neutrons in the case of silver, arsenic, antimony and mercury

The curves in the latter case were always of the same general type consisting of an initial component rapidly decreasing with boron absorber thickness superimposed on a component varying very slowly with boron thickness It is shown that the initial component is to be associated with a nuclear resonance level corresponding to very small neutron energy, while the other component is due to neutrons captured into resonance levels of greater energy The curves are interpreted in terms of the Bethe Placzek formulation of Bohr's theory of nuclei

In the case of silver, absorption measurements in boron have in addition been made of those neutrons which excite the γ radiations of 22 and 138 sec half period

The question of whether a nucleus which on neutron capture can emit both and γ radiation will have the same resonance levels corresponding to both methods of reorganization is discussed For silver the evidence points to the existence of separate sets of energy levels for γ ray emission and γ decay but it is not sufficiently conclusive to settle finally this point

---

The development in vitro of the mammalian gonad Ovary and ovogenesis By P N MARTINOVITCH (*Communicated by F H A Marshall, F R S —Received* 15 *December* 1937)

Ovaries of young and embryonic rats and mice were successfully cultivated *in vitro* and a progressive differentiation of the germ cells obtained The tissue was grown by the watch glass method, the medium consisting of equal parts of fowl plasma and fowl embryo extract In each case, one gonad was fixed as a control, whilst the other was cultivated *in vitro*

If the ovaries are explanted at the 15th–16th day of embryonic development when primitive germ cells are only present and ovogenesis has not yet begun, rapid division of the oogonia followed by the typical synaptene stages of meiosis proceed uninterruptedly in the normal way The diplotene phase of meiosis is succeeded by a resting stage when the individuality of the chromosomes becomes indistinguishable Most of the resting oocytes remain unchanged until the end of the culture period (about 1 month), but some enlarge and after about 18 days' cultivation attain full size Thus a primitive germ cell may develop *in vitro* from an oogonium into a full-size ovum In the explants growth of the ovum is not always accompanied by growth of the follicle and an ovum may reach its full size in the complete absence of a follicle New germ cells are not formed *in vitro* except by the mitotic division of the original oogonia

The ovaries of older embryos and post embryonic animals also develop readily *in vitro*

The influence of wall oscillations, wall rotation, and entry eddies on the breakdown of laminar flow in an annular pipe By A FAGE (*Communicated by G I Taylor, F R S —Received* 15 *December* 1937)

(1) *Scope of work* Experiments have been made to determine the effects of disturbances of known character on the laminar flow of water in a long pipe of annular cross section The disturbances considered are those due to axial oscillations of the inner wall of the pipe, to oscillations of the inner wall about its axis, and to both weak and intense entry eddies Experiments on the breakdown due to a uniform rotation of the inner wall (outer wall fixed) have also been made

The breakdown of flow near a plane surface oscillating in a stationary fluid has been observed

Theoretical relations for the flow of a viscous fluid through an annular pipe, under the influence of a pressure gradient parallel to the axis, with the inner wall oscillating axially and the outer wall fixed have been obtained

(2) *Conclusions* The frequency of the axial oscillation of the inner wall when a departure from laminar flow occurs depends on the axial amplitude of the wall and the viscosity of the fluid, and is independent, within the accuracy of measurement, of the velocity of axial flow The Reynolds number of disturbance, defined as the product of the velocity amplitude at the wall and a length $2\pi\sqrt{\nu/\pi f}$ (where f = frequency) divided by the viscosity, at which a departure from laminar flow occurs does not change appreciably over a wide range of amplitude

The results with the inner wall of the pipe oscillating about its axis suggest that the flow remains laminar up to the critical Reynolds number of disturbance measured with the inner wall oscillating axially

Visual observation suggests the presence of rotating bands of fluid with their axes parallel to the direction of oscillation at the breakdown

When the inner wall rotates at a uniform speed (outer wall stationary), the critical rotational speed increases with the axial speed of flow, and the critical number for no axial flow, predicted by extrapolation of the curve drawn through the numbers measured with axial flow, is in close agreement with Taylor's theoretical number

It is shown that the early breakdown of laminar flow associated with intense entry disturbances can be caused by very weak entry disturbances, provided they are in the form of discrete eddies

---

The formation of mercury molecules II By F L ARNOT and M B McEWEN (*Communicated by O W Richardson, F R S —Received* 17 *December* 1937)

An investigation of the formation of ionized mercury molecules has been made by the balanced space charge method The results obtained in this work, together with those previously obtained with a mass spectrograph by Arnot and Milligan, show that molecular ions are formed by the three processes

$$\mathrm{Hg}'(2\,{}^3\mathrm{P}_1) + \mathrm{Hg}'(2\,{}^3\mathrm{P}_0) \rightarrow \mathrm{Hg}_2^+ + \epsilon \qquad (1)$$

$$\mathrm{Hg}'(4\,{}^1\mathrm{P}_1) + \mathrm{Hg} \rightarrow \mathrm{Hg}_2^+ + \epsilon \qquad (2)$$

$$\mathrm{Hg}'(5\,{}^3\mathrm{P}) + \mathrm{Hg} \rightarrow \mathrm{Hg}_2^+ + \epsilon \qquad (3)$$

Ions formed by each of these processes have been detected The appearance potentials for these three processes are respectively 4 86 and 9 722 V for (1) and (2) and 9 792, 9 796, 9 817 V for the triplet process (3) Ions formed by process (1) which requires a collision between two excited atoms were detected at a vapour pressure of 0 1 mm of Hg at 0° C Ions formed by processes (2) and (3) which involve a collision between one excited atom and a normal atom were detected at pressures of the order of 0 01 mm

The ionization potential of the mercury molecule lies between $9\ 52+D$ and $9\ 32+D$, where D is the work of dissociation of the normal molecule Using Winan's value for D the limits are 9 67 and 9 47 V

Excited atoms in states other than P states do not apparently form ionized molecules by attachment even though they have more than sufficient energy to do so

---

On the occurrence of helium in beryls By J W J Fay, E Glückauf and F A Paneth (*Communicated by J C Philip, F R S —Received 17 December 1937*)

Various specimens of old beryllium metal have been analysed for helium As the method employed, in spite of its sensitivity, failed to detect any traces of helium it must be concluded that the spontaneous production of helium in beryllium is less than $1\ 3\times10^{-11}$ c c of helium per gram beryllium per year

From this figure it follows that the helium content of beryls cannot be explained as a consequence of the spontaneous disintegration of a beryllium isotope of mass eight Even the assumption that such a beryllium isotope has been present in previous geological periods and now mostly decayed is not compatible with the very low limit found for present day helium production

In recent years the helium content of beryls has been attributed to the influence of γ rays from radioactive minerals in the neighbourhood of the beryls, and to cosmic radiation From the figures obtained by measurement of the amount of helium produced in beryllium by γ rays it follows that the influence of the natural sources of γ radiation is not nearly sufficient to explain the helium content of beryls

Since, therefore, beryllium does not produce adequate amounts of helium, either under the influence of external radiation or as a consequence of spontaneous disintegration, it seems that the helium content of beryls is not connected with its beryllium content at all, but is due to some other chemical element

---

Identification and measurement of helium formed in beryllium by γ-rays. By E Glückauf and F A Paneth (*Communicated by J C Philip, F R S —Received 17 December 1937*)

While it has been known for some time that the beryllium nucleus, irradiated by γ rays, emits neutrons, it could not be decided whether the nucleus is thereby transformed into a stable isotope of beryllium of mass eight, or into two helium atoms

By a micro-chemical method helium was detected in beryllium after irradiation by the γ rays of radon In one of the experiments the number of neutrons simultaneously emitted during the irradiation was determined by measuring the helium produced by these neutrons in methyl borate, a comparison of the two helium quantities showed that the main final product of the γ irradiation of beryllium is helium and not the beryllium isotope

---

The nature of the penetrating component of cosmic rays By P M S BLACKETT, F R S

(*Received* 18 *December* 1937)

Further measurements of the energy loss of cosmic rays in metal plates have confirmed the former result of Blackett and Wilson that nearly all the rays with energy less than 2×10^8 e volts are electronic in character For higher energies the mean loss is much smaller A reinterpretation of the data for these higher energies in the light of the cascade theory of showers has confirmed the conclusion of Neddermeyer and Anderson that there must be some energetic but absorbable rays which can be identified with fully radiating electrons But the number of these rays is quite small, not more than 1 % of the rays at sea level with energy over 3×10^8 e volts being electrons

Both direct and indirect evidence is found which shows that the energetic penetrating rays actually become, as they slow down, the absorbable rays of low energy which are indistinguishable from electrons The main requirement of a theory of the penetrating component is to explain this striking property The two most obvious types of explanation are as follows The rays can be considered as heavy when energetic, but to change their mass suddenly when their energy falls below the critical energy Or, the rays are assumed to have electronic rest mass, but to be distinguished from normal electrons by some new property, which has the effect of making their energy loss vary with their energy and the nature of the absorber

---

Oscillography of adsorption phenomena 3 Rates of deposition of oxygen upon tungsten By M C JOHNSON and A F HENSON

(*Communicated by A M Tyndall, F R S —Received* 18 *December* 1937)

An experiment is devised for investigating the passage of small and reproducible quantities of oxygen over a hot tungsten filament whose surface purity can be ascertained by observation of its thermionic emission immediately before each admission of gas The oxygen pressure at the filament is made to rise within a fraction of a second from a high vacuum to a transient value with a flat maximum of 10^{-3} to 10^{-4} mm This enables the rate of primary adsorption to be isolated, and its dependence on gas pressure and on temperature to be determined, without the resulting monolayer becoming sufficiently dense to allow oxidation and appreciable thinning of the filament The adsorption is traced by photographing the accompanying fall of electron emission by means of the oscillographic technique described in previous papers Within a range of temperature and pressure such that evaporation is

negligible and interaction between adsorbed particles is rare, the uncovered fraction of the surface decreases exponentially with time over a portion of each experiment This enables a condensation constant to be obtained and its temperature dependence evaluated The temperature coefficient which is found implies a heat of activation of 24,000 ± 3000 cal per mol in the vicinity of 2200° abs The bearing on theories of activated adsorption is briefly discussed

---

The law of error and the combination of observations By H JEFFREYS, F R S (*Received* 18 *December* 1937)

The limitations of the theoretical grounds for the normal law of errors of observation are discussed, and seven series of observations capable of providing tests of the law are examined It is found that the χ^2 test, as usually employed, is not sufficiently sensitive to establish departures from the normal law A wider grouping however, reduces the random error of χ^2 enough to show them clearly, though it is still less sensitive than the ratio of the maximum likelihood solution for the departure to its standard error It appears that no form of the test is of much use when the law to be tested implies very small expectations in some of the groups

An approximation to the method of maximum likelihood for Pearson laws of types II and VII is developed, and extensions to types I and IV are suggested The approximation does not require excessive labour or the retention of a large number of figures It is found that the various series of data give laws ranging from type II with index 4 5 to type VII with index 4 3, and that the index is closely correlated with the degree of correlation of the errors within groups of successive observations An extrapolation using this correlation suggests that genuinely independent observations would follow a type VII law with index between 3 and 4 Methods of combining observations derived from such a law and determining their uncertainties are provided It appears that a number of discrepancies in physics and astronomy that have been accepted as systematic may turn out to be random, since with such a law large random errors may occur more often than with the normal law, if the mean and the mean square deviation are still used as estimates

---

The crystalline structure of steel at fracture By H J GOUGH, F R S and W A WOOD (*Received* 20 *December* 1937)

In a previous investigation, precise methods of X ray diffraction were employed to study the effects on the crystalline structure of normalized mild steel of deformation and fracture under static and fatigue stresses It was established that failure by static and cyclic stressing was characterized by exactly the same kind of progressive deterioration of the original crystals leading to a completely fragmented system of crystallites of random orientation In addition to this state, it was considered that the presence of marked lattice distortion in the crystallites was also a necessary condition for fracture but direct and indisputable evidence on this aspect could not be obtained under the conditions of the previous experiments The further work now described was, therefore, undertaken A mild steel *initially* in a heavily cold worked condition has been used, enabling the process of further deformation and fracture to be studied the fatigue method of stressing has been employed in view of the unique advantages offered by it

While it is found that under safe ranges of stress, no measurable change occurs in the structure, cycles of an unsafe range produce a progressive effect on the efficiency of scattering of the X rays only to be accounted for by the incidence of a heavy lattice distortion in the crystallites forming the fragmented material The establishment of this condition of severe internal stress at the fracture stage is the most important result of the work In addition, however, informative work has also been done on *normalized* mild steel using an improved X ray technique among other results, it is shown that the rate of modification of structure, in a fatigue test, decreases as the test proceeds tending to the establishment of a stable structural condition

---

Studies on chromatic behaviour in crustacea The receptive mechanism of the background response By H G Smith (*Communicated by L Hogben, F R S —Received 21 December 1937*)

The pigmentary effector system of *Ligia* consists of melanophores and xanthophores separate in some regions and in other regions superimposed so as to form complex chromatophores The melanophores are the chief agents of visible colour change and their response to illumination is partly direct and partly controlled by the eyes The primary response is of the type described for decapods and mysids by Keeble and Gamble and for vertebrates by Laurens and Hogben Intense illumination increases the diffusion of the pigment The secondary or visual response to background is also of the type described for *Macromysis* and vertebrates Overhead illumination in a black container makes the animal dark, and overhead illumination of an animal in a white container makes it pale Blinded animals are not so dark as seeing animals kept on a black background In darkness they are still more pale owing to elimination of the primary response The time relations of the background response show that it is controlled by two hormones The existence of a "supernormal phase" in the transition from white background equilibrium reinforces this conclusion So also do experiments which show how the supernormal phase may be accentuated or eliminated The effect of varying (*a*) the direction of incident illumination, (*b*) limiting the light scattered on to the eye make it possible to distinguish two groups of ocelli One (dorsal) is accessible to direct overhead illumination and is responsible for initiating the discharge of the hormone which evokes melanophore expansion The other (latero ventral) picks up light scattered from the immediate surroundings and is responsible for initiating the discharge of the hormone which evokes melanophore contraction The effects of painting the whole eye, the ocelli of the upper half only, or the ocelli of the lower half alone, directly confirm this conclusion The blue end of the spectrum is most effective for evoking the visual or secondary response

---

On the chromatic behaviour of elasmobranchs

By H WARING (*Communicated by L Hogben, F R S —Received* 21 *December* 1937)

With few exceptions, all Elasmobranchs so far investigated pale on a white back ground and darken on a black one These changes are brought about by the contraction and expansion of dermal and epidermal melanophores The expanded condition is due to a blood circulated hormone B produced by the neuro intermediate lobe of the pituitary Parker and his associates believe that pallor is brought about by the stimulation of melanophoric nerves The present paper describes two classes of experiments which show that the anterior lobe produces a hormone which antagonizes the B substance produced by the neuro intermediate lobe These are (*a*) effects of removal of the anterior lobe, (*b*) differential tolerance to B containing extracts of normal animals, animals from which the whole gland has been removed, animals deprived of the neuro intermediate lobe alone Reasons are advanced for believing that the class of experiments upon which Parker's interpretation is based is incapable of yielding unequivocal evidence in favour of it

---

On the neutrino theory of light

By M H L PRYCE (*Communicated by P A M Dirac, F R S —Received* 21 *December* 1937)

This paper brings to light a grave difficulty for the neutrino theory of light Starting from assumptions about the neutrino sufficiently general to include the models which have been studied by Jordan Kronig and others (with the exception of Scherzer's attempt, which is not strictly a neutrino theory), and working with the amplitudes of the second quantization as the most suitable mathematical apparatus, one sets up the most general theory consistent with Jordan's hypothesis The conditions under which this will lead to a satisfactory theory of light are (1) that certain commutation rules be satisfied, (2) that the theory be invariant under a change of co ordinate system In order to study the second of these it has been necessary to analyse rather carefully the transformation of the amplitudes under certain types of rotation and this reveals an arbitrariness in the choice of certain phases A condition for the invariance of the theory is that the results be independent of the way in which these phases are chosen

From this point onward a straightforward analysis leads to the result that the conditions cannot be satisfied simultaneously The invariance requires that the neutrino which interacts with the atom should reverse its spin, a result which could also be derived from considerations of the conservation of angular momentum, and an essentially simple though rather tedious calculation shows this to be inconsistent with the commutation rules

The introduction gives an account of the aims of the neutrino theory of light, the problems which it meets and the attempts that have been made to solve them

The absorption spectrum of sulphur dioxide and carbon dioxide in the vacuum ultraviolet

By W C PRICE and D M SIMPSON (*Communicated by O W Richardson, F R S —Received* 23 *December* 1937)

The absorption spectra of sulphur dioxide and carbon disulphide have been investigated by means of a vacuum spectrograph down to 1000 A For both molecules the systems of bands can be divided into two classes (1) those which exhibit wide vibrational structure, (2) those which exhibit little or no vibrational structure The former class probably correspond to transitions to anti bonding molecular orbitals, while the latter are due to the transitions of comparatively non bonding electrons to excited orbitals, which are mainly atomic in character In the case of sulphur dioxide, the extrapolation of the bands of class (2) to their limit gives a value of 12 05 ± 0 05 V for the ionization potential of the molecule A similar procedure for carbon disulphide yields the much more accurate values 10 083 and 10 027 V for ionization to the two components of the doublet state $^2\Pi_g$ of CS_2^+ ($\Delta\nu = 436$ cm $^{-1}$) The experimental evidence indicates that while carbon disulphide is slightly bent in the earlier stages of the excitation, it finally returns to a linear configuration in CS_2^+ A vibrational analysis of the bands of class (1) is also given, and some general features of the electronic spectra of polyatomic molecules are discussed

---

Significance tests when several degrees of freedom arise simultaneously

By H JEFFREYS, F R S (*Received* 29 *December* 1937)

Tests are provided for the significance of an estimated departure from a uniform distribution of chance, and of the coefficients of new functions introduced into an empirical law designed to fit a series of measures, in each case where several degrees of freedom may be expected to arise together if one of them does A test is also given for the independence of errors of observation when the means of groups of consecutive observations are compared with the standard deviation of the entire set Applications are made to the secular perturbations of the inner planets and to Pearson's data for errors of observation

---

Vacuum wavelength measurements in the iron spectrum by means of the reflexion echelon grating

By W E WILLIAMS and A MIDDLETON (*Communicated by O W Richardson, F R S —Received* 29 *December* 1937)

The vacuum wave lengths of forty seven iron lines have been determined directly against the red cadmium standard by means of a new method employing a reflexion echelon in vacuum so that no corrections for the refractive index of normal air need be made By using a Schüler type hollow cathode source, much narrower lines are obtained than is possible with an iron arc and the concordance of independent determinations justify the values being given to eight significant figures The work is being extended further to the ultra violet

The forces on a cylinder, or flat plate, submerged in a uniform stream

By T H HAVELOCK, F R S (*Received* 29 *December* 1937)

General expressions are obtained for the resistance, lift and moment for a cylinder submerged in a uniform stream, including circulation, the analysis being an extension of that used in a previous paper on the circular cylinder Application is made to the elliptic cylinder, the solution involving an integral equation which may be solved by successive substitution to any required degree of approximation

Special consideration is given to the flat plate, which is treated as the limiting case of an elliptic cylinder, expansions are obtained for the resistance, lift and moment, and graphs of the two latter quantities are given

The general expressions include, as special cases when gravity is neglected, the problem of a flat plate in a uniform stream bounded by a free surface or by a plane wall, these cases are examined with reference to the interference of the surface of the sea or of the ground upon the lift and moment of an aerofoil

---

Collective electron ferromagnetism

By E C STONER (*Communicated by O W Richardson, F R S —Received* 29 *December* 1937)

General equations are obtained, using Fermi Dirac statistics, for the magnetic moment, M, of a number, N, of electrons in an unfilled energy band, for which the interchange interaction effects give rise to a term in the energy expression proportional to the square of the magnetization The relative magnetization, $\zeta (= M/N\mu)$, is expressible as an implicit function of the reduced field, $\mu H \epsilon_0$, temperature, kT/ϵ_0, and interaction energy coefficient, $k\theta'/\epsilon_0$, where ϵ_0 is the Fermi energy without interaction Particular limits correspond to equations for electron spin paramagnetism, and to the Weiss Heisenberg "classical" equations for ferro and paramagnetism

The equations are solved numerically, using tables of Fermi Dirac functions (McDougall and Stoner), in such a way as to give kT/ϵ_0 as a function of ζ for a range of values of $k\theta'/\epsilon_0$ The results are shown in a series of tables and diagrams

The character of the dependence of ζ on kT/ϵ_0, or on T/θ, depends on the ratio $k\theta'/\epsilon_0$ A necessary condition for ferromagnetism is $k\theta'/\epsilon_0 > \frac{2}{3}$, while for $k\theta'/\epsilon_0 < 2^{-\frac{1}{3}}$ ($= 0$ 793701), the relative magnetization at absolute zero, ζ_0, is less than unity For small values of ζ_0, the magnetization temperature curve is closely represented by

$$(\zeta/\zeta_0)^2 = 1 - (T/\theta)^2,$$

but the curve does not change monotonically to the classical form as $k\theta'/\epsilon_0$ increases

Series expansions appropriate for particular ranges are given for $k\theta'/\epsilon_0$ as a function of $k\theta/\epsilon_0$, and of kT/ϵ_0 and ζ, and expressions are derived for the variation of the magnetization near the Curie point and at low temperatures

A full discussion of the experimental results is deferred, but it is shown that the lack of agreement between the values deduced for the saturation moment from the paramagnetism above the Curie point and from the low temperature magnetization receives an immediate interpretation

A new basis for cosmology By P A M DIRAC, F R S (*Received* 29 *December* 1937)

It is proposed that all the very large dimensionless numbers which can be constructed from the important natural constants of cosmology and atomic theory are connected by simple mathematical relations involving coefficients of the order of magnitude unity The main consequences of this assumption are investigated and it is found that a satisfactory theory of cosmology can be built up from it

---

The polarization of a calomel electrode By LORD ROTHSCHILD (*Communicated by J Gray, F R S —Received* 30 *December* 1937)

The maximum current density which an anodic calomel electrode will sustain over reasonable periods of time (5 hr) without polarization is 15 amp $^{-6}$ cm $^{-2}$

Cytochrome oxidase By D KEILIN, F R S and E F HARTREE (*Received* 30 *December* 1937)

1 The addition of 10^{-5} to 10^{-4} M cytochrome c to heart muscle oxidase preparation greatly increases the rate of catalytic oxidation of a number of diamines and polyphenols such as p phenylenediamine, p aminophenol, hydroquinone, adrenaline, catechol, etc

2 All these compounds readily reduce a solution of oxidized cytochrome c

3 The rate of oxidation of these compounds by oxidase preparation may increase, on addition of cytochrome c, up to 30–40 times

4 The much slower oxidation of these compounds without addition of cytochrome c is due to cytochrome already present in the oxidase preparation

5 The oxidation of all these compounds therefore is not catalysed directly by indophenol oxidase but through co operation with cytochrome In other words, the indophenol reaction of cells and their extracts takes place through cytochrome

6 The only catalytic property which can be definitely ascribed to the muscle oxidase (known as indophenol oxidase) is the oxidation of reduced cytochrome The correct name for this enzyme should, therefore, be *cytochrome oxidase*

7 The explanations as to the nature, distribution and catalytic activity of this enzyme proposed by Remesow, Bigwood, Harrison and Shibata with his co workers, are discussed and shown to be incorrect

8 There is no necessity to assume the existence in the respiratory system of the cell of a component undergoing reversible "oxygenation" as was proposed by Shibata and his co workers

Relation between the uterus and the ovaries in the pregnant hamster.

By M Klein (*Communicated by A S Parkes, F R S —Received 3 January* 1938)

1 Experiments on the physiology of pregnancy have been carried out on the golden hamster

2 Complete ovariectomy performed between the 9th and the 13th day results in the termination of pregnancy

3 The injection of progesterone alone failed to maintain pregnancy and the typical vaginal mucification in hamsters ovariectomized about the 10th day after mating The simultaneous injection of progesterone and oestrone did, however, maintain pregnancy and vaginal mucification in ovariectomized females, and in several cases we have obtained parturition at term with living litters These results suggest that, in the hamster, progesterone and oestrone may exercise endocrine control of pregnancy

4 Complete removal of the pregnant uterus between the 8th and the 13th day causes the premature and rapid regression of the corpora lutea of pregnancy, and the ovarian cycle restarts When the foetuses only are removed, by caesarian section, and the placentae remain inserted, the corpora lutea of pregnancy are maintained and inhibition of the ovarian cycle continues

5 These results agree with and complete our previous experiments on the rabbit and the rat They show that there is a close correlation between the pregnant uterus and the ovaries, though we cannot yet postulate a satisfactory interpretation of the mechanism of the connexion

---

Progressive lightning V A comparison of photographic and electrical studies of the discharge process

By B F J Schonland, D B Hodges and H Collens (*Communicated by Sir Charles Boys, F R S —Received 5 January* 1938)

A detailed comparison is made of the results of Boys' camera observations of the luminosity changes during the lightning discharge and of observations with a cathode ray oscillograph of the electric field In certain cases these observations were simultaneous

The polarity of the discharges to ground was found to be negative in all cases examined

The results confirm and extend the identifications of the electric field changes made by Appleton and Chapman During a ground stroke these begin with the slow lowering along the leader channel of about 60 % of the cloud charge tapped, and are followed by the sudden removal of this leader charge to ground by the return streamer and by the slower passage of the remainder of the cloud charge to ground

The more important transient induction and radiation effects are associated with the stepped leader to first strokes and with the start of the return streamer

Since the return streamer is absent in most discharges within the cloud it is suggested that discharges to ground are the chief cause of atmospherics from distant thunderclouds

A suggestion for unifying quantum theory and relativity By M Born (*Communicated by E T Whittaker, F R S —Received 5 January 1938*)

The fact that the fundamental laws of quantum mechanics are symmetrical in space time x^k and momentum energy p_k can be generalized to a "principle of reciprocity" according to which the x space and the p space are subject to geometrical laws of the same structure, namely a Riemannian metric In analogy with Einstein's closed x world one has to assume that energetically closed systems (as elementary particles, nuclei) must be described by help of a hyperspherical p space A consequence of this assumption is a modification of the formula for the number of quantum states in an element of the p space The application of this formula to quantum electrodynamics leads to a finite zero energy of the vacuum, a finite self energy of the electron, etc Deviations from Planck's law and the Stephan Boltzmann law of radiation, and the caloric properties of gases are predicted for very high temperatures

---

X-ray studies of the structure of hair, wool and related fibres IV The molecular structure and elastic properties of the biological cells By H J Woods (*Communicated by Sir William Bragg, O M, P R S —Received 5 January 1938*)

From a study of the cells isolated from various types of animal hair it is concluded that they contain crystalline keratin, and when the cells are long, tissues can be built up to give X ray photographs closely comparable with those of the fibres Cells obtained by retting fibres previously set in steam are found to be stretched correspondingly, although their measured lengths after retting are somewhat shorter than would be expected This lag is associated more with the early stages of extension, indicating that the elastic phase K_1 extends inside the cells, as well as between them The X ray photograph of elongated cells from fibres set at high extensions is the ordinary one of β keratin The elongated cells only recover partially when they are boiled in water, and the permanent set remaining is of the same order as that of fibres steamed for the same time at the same extension as those from which the cells were obtained Permanent set is thus also a property of the cells themselves

Cotswold wool fibres which have been relaxed in dilute caustic soda can also be disintegrated into cells, which recover during the process to rather less than the normal length These relaxed cells can be supercontracted still further by boiling water In this way cells of about half the normal length may be obtained, their crystalline part shows itself in the β form, just as does that of highly supercontracted fibres

By the combined action of lateral pressure and steam the cells from normal fibres can be compounded into coherent transparent sheets which are elastic in cold water for extensions up to about 50 % and extensible in steam or caustic soda by twice this amount These sheets can be relaxed, set, and supercontracted, in the same way as fibres, and their tensile strength in water is about 25 % of that of the fibres It is suggested that the development of coherence during the treatment in steam is of a similar type to the development of permanent set in the stretched fibre

Since such differences as exist between the properties of the cells and those of the fibres are only of the same order as those which occur between various types of hairs, it is clear that the behaviour of the fibres is for the most part due to the cells The histological heterogeneity of hairs can affect only the details of their behaviour

Keratin, stretched denatured edestin, and stretched myosin are all anisotropic in their electrical properties, having different dielectric constants for fields parallel and perpendicular to the protein chains By hanging discs of these proteins in an alternating electric field it is therefore possible to determine the direction of the chains

---

Polyphenol oxidase Purification, nature and properties

By D KEILIN, F R S and T MANN (*Received* 11 *January* 1938)

1 A comparatively simple method of preparation of highly purified polyphenol oxidase from cultivated mushrooms (*Agaricus campestris*) is described

2 The activity of the enzyme in different preparations was determined by two methods (1) manometric, from the initial rate of oxygen uptake by catechol at 20° C , calculated per mg dry weight per hour and expressed as Q_{O_2}, and (2) colorimetric, from milligrams of purpurogalline formed from pyrogallol in 5 min at 20° C per mg dry weight and expressed as purpurogalline number (P N) The amount of enzyme which, under these conditions produces 1000 mg purpurogalline is taken as one enzyme unit (E U)

3 The purest enzyme preparation obtained was completely free from haematin and contained only traces of iron and manganese

4 The polyphenol oxidase of mushrooms was found to be a copper protein compound the copper being a constituent of the active group of the enzyme This extends the discovery of Kubowitz, who found that the corresponding oxidase in potatoes is a compound of copper with protein

5 No proportionality between the copper content and the enzyme activity was found for preparations ranging from P N = 0 45 to P N = 250 These preparations contain large amounts of copper which does not belong to the enzyme Within this range the amount of copper per enzyme unit (E U) gradually diminishes from 196γ to 7γ Only when the copper content reaches a low and constant value of about $3\ 2$–$3\ 5\gamma$ per E U does the copper content become strictly proportional to the enzyme activity

6 The purest enzyme preparation obtained has P N = 940, Q_{O_2} = 1,160,000 and copper content of 0 30 % As this copper content is higher than that of pure crystalline haemocyanine, which may vary according to origin, from 0 173 to 0 26 % (Hernler and Philippi), the enzyme may be considered as pure 1γ Cu of this enzyme transfers to catechol, at 20° C in 1 min , about 6000 c.mm O_2 9 6 mg of this enzyme (dry weight) was obtained from 15 kg fresh mushrooms

7 The crude extract of mushrooms readily oxidises only such compounds as catechol, pyrogallol and *p*-cresol Other related compounds such as hydroquinone and *p* phenylenediamine are hardly oxidised by this enzyme

On purification, the enzyme gradually loses the property of catalysing the oxidation of monophenols

8 All the oxidation reactions catalysed directly or indirectly by the polyphenol oxidase are inhibited by KCN, H_2S and CO

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form. These abstracts will be issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings"

---

4 FEBRUARY 1938

The influence of radiation on ionization equilibrium

By B SRIVASTARA

(*Communicated by M N Saha, F R S —Received* 11 *January* 1938)

The generalized ionization formula in the case of thermodynamical equilibrium for an assembly consisting of atoms, ions, electrons and radiation is derived, the particles being supposed to obey any statistics

The ionization produced in matter has been found for the case when the reaction space is traversed by radiation at a higher temperature The electrons have been assumed to obey F D statistics, while the atom and the ion are taken to obey Boltzmann (classical) statistics The general ionization formula for this case is derived and two limiting cases discussed The Woltjer Milne formula for ionization in any radiation field is readily obtained as a special case of the general formula The ionization formula for the case when the electrons behave as a degenerate gas is also deduced from the general formula On substituting the usual expression for the probability of ionization, viz $\psi = C\nu^{-3}$, it is found that the ionization becomes much less when the electrons are highly degenerate

Uranium Z and the problem of nuclear isomerism.

By N Feather and E Bretscher (*Communicated by E V Appleton, F R S —Received 13 January* 1938)

The radiations from uranium Z have been examined by the absorption method using a tube counter The effective quantum energy of the γ radiation is 0 70 ± 0 05 $\times 10^6$ e volts and the intensity 1 50 ± 0 25 quanta per disintegration An analysis of the β radiation into continuous spectra with limiting energies 0 56 and 1 55 $\times 10^6$ e volts, and intensities in the ratio 17 1, is suggested, though it is pointed out that the component of lower energy is probably itself complex The uranium X_2 uranium Z branching ratio has been determined as 665 ± 65 1

On the basis of these results the isomerism of the nuclei UX_2 and UZ is discussed in the light of v Weizsäcker's hypothesis, and a level scheme is put forward which appears to account for all the facts Reasons are given in support of the conclusion that uranium Z is formed from uranium X_2 in a β γ branching, rather than from uranium X_1, directly, in a β β transformation

---

On the primary decomposition of ethane and the reaction between ethane and nitric oxide

By T J Gray, M W Travers, F R S and F T White (*Received* 14 *January* 1938)

The thermal decomposition of pure ethane, and of a mixture of ethane and nitric oxide, have been studied at 590° in silica apparatus by the method of detailed analyses

The observation of Staveley (1937) that the addition of nitric oxide to ethane reduces the rate of primary decomposition is confirmed However, it is shown that a rapid reaction takes place between ethane and nitric oxide, without apparent increase in volume, which accounts for the fact that it was not observed by the pressure measurement method which Staveley employed The product of the ethane nitric oxide reaction at first increases in concentration in the gas and then diminishes It has not been isolated

The ethane nitric oxide reaction itself slows down after an interval, but as it does so a reaction which appears to involve ethane and the product of the ethane nitric oxide reaction develops One product of this reaction appears to be ethylene

The theory of free radicals cannot be used to explain the facts now put forward

It is suggested that the ethane nitric oxide reaction operating with a lower activation energy than the primary decomposition process, removes *hot* molecules at a rate faster than they can be supplied to maintain the Maxwell Boltzmann equilibrium The result is that the former process is to a considerable extent *substituted* for the latter The ethane nitric oxide reaction in turn slows down at a rate faster than can be accounted for by the removal of ethane or nitric oxide This is associated with the operation of a reaction between ethane and a product of the ethane nitric oxide reaction at an activation energy still lower than the ethane-nitric oxide reaction itself.

A method of photographing airscrew sound-waves. By W F HILTON (*Communicated by L Bairstow, F R S —Received* 15 *January* 1938)

An experimental technique has been developed for photographing the shock waves generated by an airscrew at all speeds above a certain critical speed A modified form of spark photography was employed, and photographs were obtained of the airscrew sound wave system at top speeds ranging from 0 83 to 1 2 of the velocity of sound

The photographs also reveal a trailing vortex from the blade tip, and an eddying wake behind the blade These vortices and eddies are more distinct at high speeds than at low, but there is no definite lower limit to the speed at which they can be photographed

A method is advanced for calculating the pressure amplitude of the shock waves from measurements of photographs, and also for calculating the speed at which a low wave can first appear in front of the blade The wave velocity has been calculated from the photographs, and is found to be slightly greater than that of small amplitude sound waves

A method of calculating the shape of the waves is examined, but no indication is given of a wave front at tip speeds less than that of sound, whereas a very definite wave front exists down to tip speeds as low as 0 85 of sound velocity

The problem of noise reduction in aircraft is discussed

The crystal structure and magnetic properties of $CuSO_4$ $5H_2O$ By K S KRISHNAN and A MOOKHERJI (*Communicated by Sir Venkata Raman, F R S —Received* 18 *January* 1938)

The crystal structure of $CuSO_4$ $5H_2O$ has been analysed by X ray methods by Beevers and Lipson On the basis of the structure proposed by them, and the theory developed by Van Vleck and others which attributes the magnetic anisotropy of a paramagnetic crystal to the asymmetry of the strong electric fields acting on the paramagnetic ions in the crystal, the axes of the magnetic ellipsoid of the crystal, and its eccentricity, are deduced theoretically

Detailed experimental data are given for the setting-directions and the magnetic anisotropy for several planes in the crystal The axes of the magnetic ellipsoid, and the orientations of the central circular sections of the ellipsoid (which give a measure of the relative magnitudes of the principal diameters of the ellipsoid), are determined from the magnetic data They agree well with the results deduced from the structure

The magnetic anisotropy of the ultimate paramagnetic anisotropic units, consisting each of a Cu^{++} ion and six negatively charged oxygen atoms surrounding it, which form a regular octohedron with one of its diagonals elongated, are calculated from the magnetic and the structural data for the crystal These paramagnetic unit groups are approximately tetragonal in symmetry, the susceptibility along the tetragonal axis being greater than for perpendicular directions by about 550×10^{-6} per g ion of Cu^{++} This anisotropy gives us a measure of the asymmetry of the electric fields acting on the Cu^{++} ions in the crystal

Diamagnetism of cadmium. By S RAMACHANDRA RAO and S SRIRAMAN (*Communicated by O W Richardson, F R S —Received* 18 *January* 1938)

Single crystals of cadmium were prepared by the method of slow cooling and the principal magnetic susceptibilities were determined by the Guoy method The influence of ferromagnetic impurities on the observed values was eliminated by determining the susceptibilities at different field strengths in accordance with Honda's method The mean values perpendicular and parallel to the hexagonal axis of the crystal were found to be −0 163 and −0 223 respectively The magnetic anisotropy was 1 368 and the mean diamagnetic susceptibility for a polycrystalline aggregate −0 183 The value −0 163 of $\chi_\perp$ is close to that of McLennan, Ruedy and Cohen but our value of $\chi_\parallel$ is lower than theirs This makes the magnetic anisotropy of cadmium much nearer that of zinc than the earlier values indicated

The influence of cold working on polycrystalline rods and single crystals of cadmium was investigated In the case of polycrystalline specimens, a small decrease was obtained On the other hand, the principal susceptibilities measured with single crystals showed no variation on cold working

The influence of small quantities of foreign matter was also a subject of investigation Lead when added in small amounts affected the principal susceptibility normal to the *c* axis but not the value parallel to this axis The principal susceptibilities did not alter as a result of adding small quantities of zinc

These results are examined in the light of the theory of paramagnetism of free electrons developed by Pauli, Landau and Peierls The valency electrons appear to contribute a paramagnetic component perpendicular to the *c* axis There is evidence for this from electrical conductivity data The results obtained with the alloys are examined in the light of these conclusions

---

Tests of statistical hypotheses which are unbiased in the limit By J NEYMAN (*Communicated by G U Yule, F R S —Received* 19 *January* 1938)

Propositions I and II show that the conception of a test which is unbiased in the limit, first introduced in connexion with the "smooth" test for goodness of fit, is applicable to a great variety of problems, where technical difficulties prohibit the application of unbiased tests of type A Propositions III and IV provide the possibility of determining approximately the power functions by calculating their limits, for fixed values of the standardized errors in the hypothesis tested and for the number of observations indefinitely increasing These theoretical results are illustrated on two examples

Owing to the relative easiness of handling the tests unbiased in the limit promise to be of great use But it is necessary to remember their limitations; a test "unbiased in the limit" may be biased for any finite number of observations, though the actual bias is probably frequently negligible We must remember further that the Propositions III and IV give us only approximate information concerning the power function and, what is more, the degree of the approximation is not known It follows that some further research in the matter is desirable

ABSTRACTS

OF PAPERS COMMUNICATED TO THE ROYAL SOCIETY OF LONDON

In accordance with a resolution of Council, summaries or abstracts of papers are to be published as soon as practicable The publication of such abstracts in no way indicates that the papers have been accepted for publication in any fuller form These abstracts will be issued for convenience with the "Proceedings of the Royal Society of London" but do not form a part of the "Proceedings"

---

18 February 1938

The probability of annihilation of positrons without emission of radiation

By H S W Massey and E H S Burhop (*Communicated by R H Fowler, F R S —Received* 19 *January* 1938)

When mutual annihilation occurs on collision of an atomic electron and a positron the energy liberated, instead of appearing as radiation, may be absorbed by a second atomic electron, resulting in its ejection from the atom In this paper the probability cross section for this radiationless annihilation process, involving the *K* electrons of a lead atom, is calculated with accuracy, account being taken of the repulsive influence of the nucleus, retardation, spin spin interaction and electron exchange Estimates are also made of the contributions from other pairs of electrons and the conclusion is finally reached that the total cross section for radiationless annihilation by lead electrons attains a maximum value of between 2 and 3×10^{-26} cm ^{2} for positrons with energy 300,000 e volts The possibility of observing the phenomenon is briefly discussed

---

The adsorption of vapours at plane surfaces of mica I

By D H Bangham and S Mosallam (*Communicated by D L Chapman, F R S —Received* 20 *January* 1938)

Measurements have been made of the quantities of benzene, carbon tetrachloride, and methyl alcohol adsorbed at a known surface area of mica at pressures ranging from 0 02 mm to near saturation The general shape of the benzene isotherm is sigmoid, convex to the adsorption axis at lower pressures, where the Langmuir monolayer is incomplete, but becoming markedly convex to this axis near saturation, where the film is multimolecular The same isothermal is obtained when the mica plates are separated by fine wires as when they are packed closely face to face Capillary condensation of liquid in incipient cleavages at the mica edges does not take place, and the thick films have properties which distinguish them from the bulk liquids The isotherms of benzene and methyl alcohol at lower pressures agree well with the theoretical equation for films of mobile molecules oriented end-on to the surface, but Langmuir's equation leads to limiting adsorption values for the monolayer which bear no relation to the lattice constants of the mica The isotherm of carbon tetrachloride shows marked discontinuities, and a determining factor in the adsorption of this substance appears to be the tendency of the molecules to cluster together

---

Polish on metals By W Cochrane (*Communicated by G P Thomson, F R S*—*Received* 20 *January* 1938)

When electron diffraction patterns consisting of diffuse rings are obtained by Thomson's reflexion method, their interpretation is uncertain since such diffuseness may be caused by the geometrical configuration of the surface of the specimen The pattern of two haloes obtained by previous workers from polished metal surfaces cannot, therefore, be held to prove the existence of an amorphous layer of metal although there is some evidence in favour of this view In order to eliminate this difficulty, a thin layer of gold was formed on a nickel base The gold film was polished, then stripped off and examined by transmitted electrons incident normal to the film In this case the pattern does not depend on the form of the surface of the film. The resulting pattern consisted of three diffuse rings It was deduced that the film consisted of very small crystal grains and that the atoms in the polish layer of a metal are arranged similarly to the atoms in a monatomic close packed liquid at a given instant After a period of 15 hr the polished gold film yielded a transmission pattern of sharp rings characteristic of polycrystalline gold, showing that crystal growth had taken place with a return to the normal crystal size

---

The spectrum of thallium chloride By H G Howell and N Coulson, (*Communicated by W E Curtis, F R S*—*Received* 22 *January* 1938)

The spectrum of TlCl has been studied in absorption and also in emission by means of a high frequency discharge Continua and diffuse bands together with a discrete band system have been found in absorption within the region 2500–3500 A The band system also occurs in emission. It is considered that the transition is $| \rightleftarrows {}^1\Sigma^+$ the levels having the following constants

Level	ν	ω_e	$x_e\omega_e$
$\vert$	31054 2	216 91	6 80
${}^1\Sigma^+$	Q	287 47	1 24

The analysis of this system given by Butkow is shown to be incorrect Predissociation occurs in the band system and is probably responsible for the band intensities being the same in both emission and absorption An attempt to identify the electronic levels of TlCl based upon the similarity of the spectrum to that of TlF has been made

No sign of the Tl_{203} isotope has been detected whereas the Cl isotope effect completely supports the proposed analysis

---

Hyperfine structure, Zeeman effect and isotope shift in the resonance lines of potassium By D A JACKSON and H KUHN (*Communicated by F A Lindemann, F R S —Received 24 January 1938*)

1 The intensity ratio of the hyperfine structure components of the resonance lines of K 39 was measured by the method of absorption in an atomic beam The value 1 45 found agrees with the value required by the spin 3/2 if allowance is made for the overlapping by the lines of the 14 times rarer isotope 41 The component of longer wave length was the stronger, showing that the nuclear magnetic moment of K 39 is positive

2 By using three atomic beams in series, of collimation 1/35, it was possible to resolve the lines of K 41 as two satellites on either side of the weak component of K 39 The ratio of the splittings of the lines of K 39 and K 41 is found to be 1 77, in good agreement with the value 1 80 found by Manley in magnetic deflexion experiments The magnetic moment of K 41 is positive, like that of K 39

The lines of K 41 have an isotope shift of + 0 0076 cm^{-1} relative to the lines of K 39 The theoretical centre of gravity shift, considering potassium as hydrogen like, is + 0·0087 cm^{-1}

3 The Zeeman Effect of the hyperfine structure of the line $4S_{\frac{1}{2}} - 4^2P_{\frac{1}{2}}$ of K 39 was investigated Each of the two π components was found to consist of four lines, proving that the nuclear spin of K 39 has the value 3/2, in agreement with magnetic deflexion experiments and the measurements of the intensity ratio of the hyperfine structure lines The observed positions of the lines are in close agreement with the positions required by the quantum theoretical formulae

---

Further investigations of the velocity of propagation of light *in vacuo* in a transverse magnetic field By C J BARNWELL and C C FARR, F R S (*Received 25 January 1938*)

An experiment designed to determine to a higher degree of precision than before the effect of a transverse magnetic field on the velocity of propagation of light, *in vacuo*, is described An interferometer method, using the Michelson arrangement, is employed, one interfering ray passing through the field in the gaps of three electro magnets and the other not A photoelectric cell, and amplifier, is used in conjunction with a galvanometer, to record band movements on a moving photographic strip, the resulting record being subsequently measured by means of a planimeter By taking the mean of the results from a number of such records, and by carrying out preliminary check experiments, it is considered that all the main causes tending to give a spurious effect greater than an assignable amount have been eliminated. The final experimental results indicate that, in a transverse magnetic field of 20,000 gauss, with residual air at a pressure of 0 05 mm , the velocity of light increases by 32 2 cm./sec , with a probable error of 21 cm /sec In view of the relatively large probable error, and for other reasons, it is considered that the result may be spurious All that can be said definitely is that in such a field the effect is less than 1 part in about 5×10^8

The scattering of cosmic ray particles in metal plates By P M S Blackett, F R S and J G Wilson (*Received* 27 *January* 1938)

Measurements have been made of the multiple scattering of cosmic rays in the following metal plates 0 33 cm lead, 1 0 cm lead, 2 0 cm copper The range of values of $H\rho$ of the tracks extended from 10^5 to 3×10^7 gauss cm , corresponding to electron energies of 3×10^7 to 9×10^9 e volts

The observed average angle θ of the multiple scattering is found to be nearly inversely proportional to the measured values of $H\rho$, and to be in fairly close agreement with the prediction of theory for rays of any mass, but with velocity nearly equal to c

At high energies, the observed values are somewhat higher than expected This discrepancy, which may be partly due to experimental error, will be further investigated

This result, that the scattering of the penetrating component is normal, while the radiation loss is much less than that expected for electrons, gives support for a heavier rest mass for the rays But this conclusion is not quite certain

---

The development of the spark discharge By T E Allibone and J M Meek (*Communicated by S R Milner, F R S —Received* 31 *January* 1938)

A general account is given of the development of the high voltage electric spark, based on the study of over 1000 photographs taken with a rotating film camera.

The electric spark is shown to consist of two principal components, a leader stroke and a main stroke—analogous to the lightning flash A leader stroke invariably starts from a positive electrode, and sometimes also starts from a negative electrode; sometimes the structure of the leader stroke is simple, sometimes it is of the "stepped" variety The leader stroke is always branched at many places, and the direction of branching is the direction of its propagation in space branching thus forms a criterion of the direction of leader-stroke development The main stroke develops in the reverse direction to that of the leader stroke and at a velocity too high to be recorded. The velocities of the positive and negative leader strokes are of the order of 10^6–10^7 cm /sec , the positive leaders being the faster Oscillograms of current and voltage support the photographic results Multiple stroke discharges have been produced and show the characteristics of the multiple lightning flashes, the first is initiated by a slow leader stroke and is branched, the subsequent main strokes are either without leader strokes or the resolution of the camera is inadequate to show them: they are not branched.